

Sushabhan Choudhury
Ranjan Mishra
Raj Gaurav Mishra
Adesh Kumar *Editors*

Intelligent Communication, Control and Devices

Proceedings of ICICCD 2018

Advances in Intelligent Systems and Computing

Volume 989

Series Editor

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

Advisory Editors

Nikhil R. Pal, Indian Statistical Institute, Kolkata, India

Rafael Bello Perez, Faculty of Mathematics, Physics and Computing,
Universidad Central de Las Villas, Santa Clara, Cuba

Emilio S. Corchado, University of Salamanca, Salamanca, Spain

Hani Hagras, School of Computer Science and Electronic Engineering,
University of Essex, Colchester, UK

László T. Kóczy, Department of Automation, Széchenyi István University,
Gyor, Hungary

Vladik Kreinovich, Department of Computer Science, University of Texas
at El Paso, El Paso, TX, USA

Chin-Teng Lin, Department of Electrical Engineering, National Chiao
Tung University, Hsinchu, Taiwan

Jie Lu, Faculty of Engineering and Information Technology,
University of Technology Sydney, Sydney, NSW, Australia

Patricia Melin, Graduate Program of Computer Science, Tijuana Institute
of Technology, Tijuana, Mexico

Nadia Nedjah, Department of Electronics Engineering, University of Rio de Janeiro,
Rio de Janeiro, Brazil

Ngoc Thanh Nguyen, Faculty of Computer Science and Management,
Wrocław University of Technology, Wrocław, Poland

Jun Wang, Department of Mechanical and Automation Engineering,
The Chinese University of Hong Kong, Shatin, Hong Kong

The series “Advances in Intelligent Systems and Computing” contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy systems, evolutionary computing and the fusion of these paradigms, social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society, cognitive science and systems, Perception and Vision, DNA and immune based systems, self-organizing and adaptive systems, e-Learning and teaching, human-centered and human-centric computing, recommender systems, intelligent control, robotics and mechatronics including human-machine teaming, knowledge-based paradigms, learning paradigms, machine ethics, intelligent data analysis, knowledge management, intelligent agents, intelligent decision making and support, intelligent network security, trust management, interactive entertainment, Web intelligence and multimedia.

The publications within “Advances in Intelligent Systems and Computing” are primarily proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

**** Indexing: The books of this series are submitted to ISI Proceedings, EI-Compendex, DBLP, SCOPUS, Google Scholar and Springerlink ****

More information about this series at <http://www.springer.com/series/11156>

Sushabhan Choudhury · Ranjan Mishra ·
Raj Gaurav Mishra · Adesh Kumar
Editors

Intelligent Communication, Control and Devices

Proceedings of ICICCD 2018



Springer

Editors

Sushabhan Choudhury
Department of Electronics
University of Petroleum and Energy Studies
Dehradun, Uttarakhand, India

Ranjan Mishra
Department of Electronics
University of Petroleum and Energy Studies
Dehradun, Uttarakhand, India

Raj Gaurav Mishra
Department of Electronics
University of Petroleum and Energy Studies
Dehradun, Uttarakhand, India

Adesh Kumar
Department of Electronics
University of Petroleum and Energy Studies
Dehradun, Uttarakhand, India

ISSN 2194-5357

ISSN 2194-5365 (electronic)

Advances in Intelligent Systems and Computing

ISBN 978-981-13-8617-6

ISBN 978-981-13-8618-3 (eBook)

<https://doi.org/10.1007/978-981-13-8618-3>

© Springer Nature Singapore Pte Ltd. 2020

This work is subject to copyright. All rights are reserved 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 Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies, Dehradun, has organized the Third International Conference on Intelligent Communication, Control and Devices (ICICCD 2018) during December 21–22, 2018. The conference focusses on the integration of intelligent communication systems, control systems and devices related to all aspects of engineering and sciences. ICICCD 2018 aims to provide an opportune forum and vibrant platform for researchers, academicians, scientists and industrial practitioners to share their original research work, findings and practical development experiences. The proceedings are published in the Advances in Intelligent Systems and Computing (AISC) book series of Springer.

The general aim of the conference is to promote international collaboration in education and research in all fields and disciplines of engineering. ICICCD 2018 is an international forum for those who wish to present their projects and innovations, having also the opportunity to discuss the main aspects and the latest results in the field of education and research.

The organizing committee is extremely grateful to the authors from India and abroad who had shown tremendous response to the call for papers. Nearly 200 papers were submitted from the researchers, academicians and students on wide areas of three parallel tracks such as intelligent communication, intelligent control and intelligent devices.

We are obliged to our Honorable Chancellor Dr. S. J. Chopra, Vice Chancellor Dr. Deependra Kr. Jha, Dean (Academic Development and Innovations) Dr. Kamal Bansal and Dean (School of Engineering) Dr. Suresh Kumar for their confidence they have invested on us for organizing ICICCD 2018.

We extend our thanks to all faculty members, external reviewers and staff members with different committees for organizing the conference and making it a grand success.

Dehradun, India

Dr. Sushabhan Choudhury
Dr. Ranjan Mishra
Mr. Raj Gaurav Mishra
Dr. Adesh Kumar

Steering Committee

Patron(s)

Dr. S. J. Chopra, Chancellor, UPES, Dehradun, India

Dr. Deependra Kumar Jha, Vice Chancellor, UPES, Dehradun, India

General Chair

Dr. Kamal Bansal, Dean, ADI, UPES, Dehradun, India

Dr. Suresh Kumar, Dean, School of Engineering, UPES, Dehradun, India

Organizing Chair

Dr. R. Gowri, HoD, Electrical and Electronics Engineering, UPES, Dehradun, India

Organizing Co-Chair

Dr. Sushabhan Choudhury, Professor, Electrical and Electronics Engineering, UPES, Dehradun, India

Dr. Ranjan Mishra, Associate Professor, Electrical and Electronics Engineering, UPES, Dehradun, India

Dr. Madhu Sharma, Associate Professor, Electrical and Electronics Engineering, UPES, Dehradun, India

Program Chair

Dr. Manish Prateek, Director, School of Computer Science, Professor, UPES, Dehradun, India

Dr. S. K. Banerjee, Associate Dean, UPES, Dehradun, India

Dr. Piyush Kuchhal, Professor, Physics Department, UPES, Dehradun, India

Prof. K. Raghwan, Electrical and Electronics Engineering, UPES, Dehradun, India

Publicity Chair

Mr. Yogesh Chandra Gupta, Electrical and Electronics Engineering, UPES, Dehradun, India

Dr. Amit Kumar Mondal, Electrical and Electronics Engineering, UPES, Dehradun, India

Public Relation Chair

Mr. Arun Dhand, Director, Media, UPES, Dehradun, India

Dr. Jitendra Kumar Pandey, Associate Dean, Research and Development, UPES, Dehradun, India

Session Management Chair

Dr. N. Prasanthi Kumari, Electrical and Electronics Engineering, UPES, Dehradun, India

Dr. Praveen Kumar Ghodke, Electrical and Electronics Engineering, UPES, Dehradun, India

Editors

Dr. Sushabhan Choudhury, Electrical and Electronics Engineering, UPES, Dehradun, India

Dr. Ranjan Mishra, Electrical and Electronics Engineering, UPES, Dehradun, India

Mr. Raj Gaurav Mishra, Electrical and Electronics Engineering, UPES, Dehradun, India

Dr. Adesh Kumar, Electrical and Electronics Engineering, UPES, Dehradun, India

Organizing Committee

Convener

Dr. R. Gowri
Dr. Sushabhan Choudhury

Co-Convener

Dr. Ranjan Mishra
Dr. Madhu Sharma

Organizing Secretary

Dr. Vivek Kaundal
Dr. Amit Kumar Mondal
Dr. Mukul Kumar Gupta

Technical Committee

Prof. K. Raghwan
Dr. N. Prasanthi Kumari
Dr. Devendra Saini
Dr. Arpit Jain
Dr. Rupendra Kumar Pachauri
Dr. S. Vamsi Krishna
Dr. Gagan Anand

Finance Committee

Mr. Vineet Mediratta
Dr. Abhinav Sharma
Mr. Roshan Kumar

Sponsorship Committee

Mr. Yogesh C. Gupta
Dr. Tarun Kumar
Mr. Manohar Reddy
Mr. Devendra Rawat
Mr. Deepak Kumar

Registration Committee

Mr. Neeraj Sharma
Mr. Deepak Kumar
Mr. Raja Rao Manda
Ms. Deepali Yadav

Hospitality and Transport Committee

Mr. Amarnath Bose
Mr. B. Khallelu Rehman
Mr. Vinay Chowdhary
Mr. Sami Rehman

Venue Arrangement and Session Management

Mr. R. M. Sharma
Ms. Monika Yadav
Ms. Isha Kansal
Mr. V. Balaji

International Committee

Dr. A. Senthil Kumar, IIRS, ISRO, Dehradun, India

Prof. A. Patnaik, IIT Roorkee

Dr. Xiao-Zhi Gao, Docent, Aalto University School of Electrical Engineering, Finland

Dr. Y. N. Singh, IIT Kanpur

Dr. Sumeet S. Aphale, University of Aberdeen, Aberdeen, UK

Dr. Tan Cher Ming, Professor, Chang Gung University, Taiwan

Dr. Vibhu Jately, KIoT, Wollo University, Ethiopia

Dr. Ashraf Darwish, Helwan University, Cairo, Egypt

Dr. Babu Sena Paul, University of Johannesburg, South Africa

Prof. Ajit Kumar Panda, Dean, NIST

Dr. Akshay Rathore, University of Victoria, Canada

Mr. Anil Kumar Rai, Principal Engineer, Samsung R&D, India

Dr. Aniruddha Mondal, NIT, Agartala, India

Dr. Ankur Saxena, Director, Surajmal College of Engineering and Management, Kicha, Uttarakhand

Dr. Arunkumar D. Mahindrakar, IIT Madras

Prof. B. K. Lande, Mukesh Patel School of Technology Management and Engineering, Mumbai, India

Dr. Debashish Ghosh, IIT Roorkee

Dr. Devendra Chack, IIT (ISM), Dhanbad, India

Dr. Durgesh Pant, Director, USAC, Dehradun, India

Dr. E. G. Rajan, Pentagram Research Centre (P) Limited, Hyderabad, India

Prof. Geetam Singh Tomar, Director, THDC, New Tehri (Uttarakhand), India

Dr. K. Giridhar, IIT Madras

Dr. J. Manuel Moreno Arostegui, Technical University of Catalunya (UPC), Barcelona, Spain

Dr. Jai Ragunathan, Marine Institute, Canada

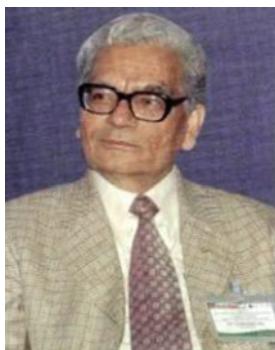
Dr. Jeevani Jayasinghe, Wayamba University of Sri Lanka

Dr. K. J. Vinoy, IISc, Bangalore, India

- Dr. Karl Heinz Steinke, Hochschule Hannover University of Applied Sciences and Arts, Germany
Mr. Krishna Kant Agrawal, Amity University, Lucknow, UP, India
Dr. Kuldeep Panwar, Delhi Technical Campus, Greater Noida, India
Dr. M. D. Singh, Thapar Institute of Engineering & Technology, Patiala, Punjab
Dr. M. V. Kartikeyan, IIT Roorkee
Dr. Manoj Duhan, Professor, DCRUST, Murthal, Haryana, India
Dr. N. S. Raghava, Professor, Delhi Technical University, New Delhi, India
Dr. Nagendra Prasad Pathak, IIT Roorkee
Dr. Nimit Gupta, Director, Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore
Dr. Nitin Khanna, IIT Gandhinagar, India
Prof. O. P. N. Calla, Director, ICRS, Jodhpur
Dr. P. ArunBabu, Marie Curie Fellow, Resiltech, Pisa, Italy
Prof. P. Chakrabarti, Director, MNNIT, Allahabad, India
Dr. Prabhat Sharma, Visvesvaraya National Institute of Technology, Nagpur, India
Dr. Preetam Suman, Integral University, Lucknow, UP, India
Dr. Pieter Jonker, Delft University of Technology, The Netherlands
Dr. Puneet Srivastava, Professor and Dean, Raj Kumar Goel Institute of Technology, Ghaziabad, India
Dr. R. C. Ramola, Dean, Himalayan School of Engineering & Technology, Dehradun, India
Dr. R. K. Chaurasia, ICFAI University, Jaipur, India
Dr. R. P. S. Gangwar, Professor, College of Technology, Pantnagar, India
Dr. Rajeev Kamal, Dayananda Sagar University, Bangalore, India
Prof. S. Jit, Professor, IIT (BHU), Varanasi-221005, India
Dr. S. N. Singh, Fellow, IEEE, IIT Kanpur
Dr. S. Sri Gowri, Professor, SRK Institute of Technology, Vijayawada, India
Dr. Safia A. Kazmi, Aligarh Muslim University, Aligarh, India
Dr. Sanjay Mathur, Professor, College of Technology, Pantnagar, India
Dr. Satya Sheel, Emeritus Professor, GLA University, Mathura, UP, India
Mr. Sérgio Ricardo Xavier da Silva, UNIFACS—Universidade Salvador, Bahia
Dr. Soumen Das, IIT Kharagpur, India
Dr. V. K. Jain, Director General, HR Institute of Technology, Ghaziabad, UP, India
Prof. Vinod Kumar, HR Institute of Technology, Ghaziabad, UP, India
Dr. Vrijendra Singh, Associate Professor, IIIT, Allahabad, India
Mr. Yashvardhan Rolyan, Project Manager, Rockwell Automation, USA
Dr. Yogesh K. Chauhan, Kamla Nehru Institute of Technology, Sultanpur, UP, India
Dr. Vipul Sharma, HoD, ECE, Faculty of Engineering & Technology—Gurukula Kangri Vishwavidyalaya, Haridwar
Dr. Vijay Nath, BIT, Mesra, Ranchi, India
Dr. Rakesh Dwivedi, Principal, CSIT, Teerthanker Mahaveer University, Moradabad, India
Dr. Ambuj Agarwal, CSIT, Teerthanker Mahaveer University, Moradabad, India
Dr. Arpit Jain, CSIT, Teerthanker Mahaveer University, Moradabad, India

Dr. Sreenivasa Rao Ijjada, GITAM University, Vishakhapatnam, India
Dr. Gaurav Verma, Jaypee Institute of Information Technology, Noida, India
Dr. Rahul Kaushik, Jaypee Institute of Information Technology, Noida, India
Dr. Sandeep Vijay, Dean, FST, ICFAI University, Dehradun, India
Dr. Gaurav Bhandari, ICFAI University, Dehradun, India
Mr. Amit Kumar, Faculty of Technology, UTU, Dehradun
Dr. Paawan Sharma, PDPU, Gandhinagar, India
Dr. Rajesh Singh, Lovely Professional University, Jalandhar, India
Dr. Anita Gehlot, Lovely Professional University, Jalandhar, India
Dr. Sonal Singhal, Shiv Nadar University, Noida, India

Keynote Speakers



Chief Guest Dr. Surendra Pal, Distinguished Scientist, ISRO, Bangalore



Prof. Raja Datta, Electronics & Electrical, Communication Engineering, IIT Kharagpur



Prof. Girish Kumar, Department of Electrical Engineering, IIT Bombay

Contents

Performance Improvement of Three-Phase Squirrel Cage Induction Motor Operating Under Rated Voltages—A Design Consideration for Rural Areas	1
Raj Kumar Saini, Devender Kumar Saini, Rajeev Gupta, Piush Verma, Neeraj Gандotra, Ashwani Sharma, Robin Thakur and R. P. Dwivedi	
A Rate Control Algorithm to Improve TCP over RFID Reader Network	13
R. Radha, Amit Kumar Tyagi and K. Kathiravan	
A Binary Harmony Search Algorithm Approach for Security Constrained Allocation of PMUs to Obtain Complete Observability	25
Manam Ravindra, Rayapudi Srinivasa Rao and Vonteddu Shanmukha NagaRaju	
Ledger-Based Sorting Algorithm	37
Nakshatra Kothi, Pradeep Laxkar, Anuj Jain and Prasun Chakrabarti	
A Survey on Wireless Optical Communication: Potential and Challenges	47
Sanjeev Kumar and Preeti Singh	
Breast Cancer Detection with a Ring-Shaped DRA (Dielectric Resonator Antenna)	53
Gagandeep Kaur and Amanpreet Kaur	
Simulation and Analysis of Distinct Apodized Profiles Using Fiber Bragg Grating for Dispersion Compensation at 100 Gbps Over 120 km	63
Ashwani Sharma, Inder Singh, Shalini Sharma, Robin Thakur and Raj Kumar Saini	

Development of Integrated Distance Authentication and Fingerprint Authorization Mechanism to Reduce Fraudulent Online Transaction	73
Vipin Khattri, Sandeep Kumar Nayak and Deepak Kumar Singh	
A Comparative Analysis of Transforms for Infrared and Visible Image Fusion	85
Apoorav Maulik Sharma, Renu Vig, Ayush Dogra, Bhawna Goyal and Sunil Agrawal	
Selection of Optimal Performance Parameters of Alumina/Water Nanofluid Flow in Ribbed Square Duct by Using AHP-TOPSIS Techniques	95
Sunil Kumar, Anil Kumar, Alok Darshan Kothiyal and Mangal Singh Bisht	
A Review: Soil Moisture Estimation Using Different Techniques	105
Jitender Pandey, Vivek Chamoli and Rishi Prakash	
Monostatic Radar Based Ultra-Wideband Microwave Imaging System Featuring a Miniature Fork Shaped Microstrip Patch Antenna with a Reduced DGS for Early Breast Tumor Detection	113
Arashpreet Kaur and Amanpreet Kaur	
An Enhanced Authentication Technique to Mitigate the Online Transaction Fraud	123
Vipin Khattri, Sandeep Kumar Nayak and Deepak Kumar Singh	
A Study of Factors to Predict At-Risk Students Based on Machine Learning Techniques	133
Anu Marwaha and Anshu Singla	
Performance Analysis of FSO Link in Log-Normal Channel Using Different Modulation Schemes	143
Harmeet Singh and Amandeep Singh Sappal	
Performance Comparison of Uncoded and ZigZag Coded IDMA	157
Shiva Kaushik and Aasheesh Shukla	
Design of High Step-up Transformerless Modular Multilevel DC-DC Converter	163
Sonal Purkait, Rahul Pandey, Lalit kumar Sahu and Saji Chacko	
Mitigation of Haze Effects on Free Space Optical Communication Using Multibeam Technique	175
Gunjan Rathore, Charu Madhu and Anaahat Dhindsa	
Spectrum Sensing in MIMO Cognitive Radio Networks Using Likelihood Ratio Tests with Unknown CSI	185
Juhি Singh and Aasheesh Shukla	

Design and Development of IoT-Based Framework for Indian Aquaculture	195
Zeenat Shareef and S. R. N. Reddy	
Top-$k\%$ Concept Stratagem for Classifying Semantic Web Services	203
Aradhana Negi and Parminder Kaur	
Performance Investigation of Fractional PI Controller in Shunt Active Filter for a Three-Phase Three-Wire System	213
Ritu Sharma, Naveen Goel, Saji Chacko and Ravi Prakash Mahobia	
High Speed 64-Bit Booth Encoded Multiplier Using Compressor	225
G. M. G. Madhuri, Ch. Aruna Kumari, Ch. Monica and N. Prasanthi Kumari	
Comparative Analysis of Least Square, Minimum Mean Square Error and KALMAN Estimator Using DWT (Discrete Wavelet Transform)-Based MIMO-OFDM System	233
Neha Awasthi and Sukesh Sharma	
An ANP-GRA-Based Evaluation Model for Security Features of IoT Systems	243
Akshay Hinduja and Manju Pandey	
A Comparative Study of 2D-to-3D Reconstruction Techniques	255
Surya Pandey	
Recent Research Trends and Methods Used in Moving Object Detection	265
Tannistha Pal, Arindam Deb, Anwesha Roy and Praseed Debbarma	
Simple Way of Tuning Centralized PI Controller for an Interacting Pilot Plant Distillation Column	275
R. Janani and I. Thirunavukkarasu	
A Hypothesis on Ideal Artificial Intelligence and Associated Wrong Implications	283
Kunal Shah, Pradeep Laxkar and Prasun Chakrabarti	
Quantum Communication Based on the Photon Area Function as Basis States for Photon Qubit	295
Vineet Kumar	
Realization of Fractional-Order Proportional Derivative Controller for a Class of Fractional-Order System in Delta Domain	303
Jaydeep Swarnakar, Prasanta Sarkar and Lairenlkpam Joypurakash Singh	

Modelling of Microwave-Based Regeneration in Composite Regeneration Emission Control System	313
Caneon Kurien, Ajay Kumar Srivastava, Karan Anand and Niranjan Gandigudi	
Image Analysis of SEM Micrograph of Co-doped ZnO-Based Oxide Semiconductors	321
Rana Mukherji, Vishal Mathur and Manishita Mukherji	
The Selection of Options for Closed-Die Forging of Complex Parts Using Computer Simulation by the Criteria of Material Savings and Minimum Forging Force	325
Volodymyr Kukhar, Elena Balalayeva, Svitlana Hurkovska, Yurii Sahirov, Oleg Markov, Andrii Prysiazhnyi and Oleksandr Anishchenko	
Optimization of DC-DC Converters for Off-Grid Lighting in Trains Using Artificial Neural Networks	333
Monu Malik and Ratna Dahiya	
A Study on Noninvasive Body Wearable Sensors	345
Shanu Bhardwaj and S. N. Panda	
An Outlier Accuracy Improvement in Shilling Attacks Using KSOM	353
Anjani Kumar Verma and Veer Sain Dixit	
Fruition of CPS and IoT in Context of Industry 4.0	367
Mishra Devesh, Agrawal Krishna Kant, Yadav Ram Suchit, Pande Tanuja and Shukla Narendra Kumar	
Leveraging Biological Dragonfly Scheme for URLLC in Industrial Wireless Network	377
Sanjay Bhardwaj, Muhammad Rusyadi Ramli and Dong-Seong Kim	
Recent Advances in Artifact Removal Techniques for EEG Signal Processing	385
Amandeep Bisht, Chamandeep Kaur and Preeti Singh	
Hierarchy of Sectors in BSE SENSEX for Optimal Equity Investments Using Fuzzy AHP	393
Kocherlakota Satya Pritam, Trilok Mathur and Shivi Agarwal	
A Novel Inset Fed Patch Antenna with Defecting Ground Plane	405
Pratik Ghosh, Kousik Roy, Atanu Nag, Rashmi Priya and Naimul Hasan	
Bionics Based Camouflage Clothing for Soldiers	415
Shweta Gupta, S. A. Hariprasad and Adesh Kumar	

An Adaptive Framework of Learner Model Using Learner Characteristics for Intelligent Tutoring Systems	425
Amit Kumar and Neelu Jyoti Ahuja	
A Study of Solar Energy in India; Utility, Status, and Procurement	435
Kamal Kashyap, S. M. Towhidul Islam Nayim, Robin Thakur, Anil Kumar and Sashank Thapa	
Necessary Precautions in Cognitive Tutoring System	445
Kevin Vora, Shashvat Shah, Harshad Harsoda, Jeel Sheth and Ankit Thakkar	
Node-Level Self-Adaptive Network Path Restructuring Technique for Internet of Things (IoT)	453
Sharma Shamneesh, Manuja Manoj and Kishore Keshav	
Design of Pole Placement-Based State Feedback Controller for a Impulse Hydro Turbine	463
Shubham Ashish, Sagar Saraswat, Sagar Garg and Alok Kumar Pandey	
Paradigm-Based Morphological Analyzer for the Gujarati Language	469
Dikshan N. Shah and Harshad Bhadka	
An Approach to Secure Collaborative Recommender System Using Artificial Intelligence, Deep Learning, and Blockchain	483
Monika Arora, Akanksha Bansal Chopra and Veer Sain Dixit	
Performance Analysis of NSL-KDD Dataset Using Classification Algorithms with Different Feature Selection Algorithms and Supervised Filter Discretization	497
Shailesh Singh Panwar and Y. P. Raiwani	
An Efficient Method for Testing Source Code by Test Case Reduction, Prioritization, and Prioritized Parallelization	513
Pradeep Udupa, S. Nithyanandam and A. Rijuvana Begum	
Hardware Realization of a Single-Phase-Modified P-Q Theory-Based Shunt Active Power Filter for Harmonic Compensation	523
Krishna Veer Singh, Ravinder Kumar, Hari Om Bansal and Dheerendra Singh	
Development and Gesture Control of Virtual Robotic Arm	531
Vineeth Veeramalla, Jagriti Mishra, C. S. Meera, Varnita Verma and Mukul Kumar Gupta	
Real Virtual IDentification (RVID): Providing a Virtual, Secure and Anonymous ID Service to Indian Users	539
Amit Kumar Tyagi, G. Rekha and N. Sreenath	

A Voting-Based Sentiment Classification Model	551
Dhara Mungra, Anjali Agrawal and Ankit Thakkar	
Dynamic Performance of Grid-Connected Wind Farms with and Without UPFC: A Case Study on Ashagoda Wind Farm	559
Asefa Sisay and Vibhu Jately	
A Study on Agent-Based Web Searching and Information Retrieval	569
Urvi Mitra and Garima Srivastava	
Impact of Internet of Things on Societal Applications	579
Tanya Srivastava and Shikha Singh	
Experimental Validation of Fully Informed Particle Swarm Optimization Tuned Multi-Loop L-PID Controllers for Stabilization of Gantry Crane System	589
Sudarshan K. Valluru, Madhusudan Singh, Daksh Dobhal, Kumar Kartikeya, Manpreet Kaur and Arnav Goel	
Design and Experimental Implementation of Multi-loop LQR, PID, and LQG Controllers for the Trajectory Tracking Control of Twin Rotor MIMO System	599
Sudarshan K. Valluru, Madhusudan Singh, Ayush and Abhiraj Dharavath	
Improving Energy Efficiency of Hybrid ARQ Scheme for Cooperative Communication in UASNs Using PSO	609
Veerapu Goutham and V. P. Harigovindan	
Implementation of FM-Based Communication System with 3-Level Parallel Multiplier Structure for Fast Transmission Using FPGA	619
Salauddin Mohammad, R. Seetha, S. Jayamangala, B. Khaleelu Rehman and D. Kumar	
Virtual Simulation of Regulated Power Supply for Various Microcontroller Boards and Their Peripherals	627
Varnita Verma, Piyush Goyal, C. S. Meera, Mukul Kumar Gupta and Piyush Chauhan	
Contribution of Learner Characteristics in the Development of Adaptive Learner Model	637
Amit Kumar and Vishal Bharti	
Event-Triggered Based Synchronization of Linear Singularly Perturbed Systems	647
Nakul Kotibhaskar, Kritika Bansal and Pankaj Mukhija	

A New Compact Multilevel Inverter Design with Less Power Electronics Component	661
Siddharth Pachpor, Subhankar Dutta, M. Jagabar Sathik and Mudit Babel	
Detection of Emergency Vehicles Using Modified Yolo Algorithm	671
Somya Goel, Abhishek Baghel, Aprajita Srivastava, Aayushi Tyagi and Preeti Nagrath	
Optimization of a Two-Cylinder Crankshaft by Computer-Aided Engineering	689
Basavaraj Talikoti, S. N. Kurbet and V. V. Kuppast	
Reducing Ping-Pong Effect in Heterogeneous Wireless Networks Using Machine Learning	697
Pragati Kene and S. L. Haridas	
Design and Analysis of Elliptic Curve Cryptography-Based Multi-round Authentication Protocols for Resource-Constrained Devices	707
Adarsh Kumar, Alok Aggarwal, Neelu Jyoti Ahuja and Ravi Singhal	
Data Mining-Based Student's Performance Evaluator	719
Ravindra Kumar, Megha Kumar and Upasna Joshi	
A Comprehensive Review on the Issues Related to the Data Security of Internet of Things (IoT) Devices	727
Shivansh Upadhyay, Shashwat Kumar, Sagnik Dutta, Ajay Kumar Srivastava, Amit Kumar Mondal and Vivek Kaundal	
SVD-Based Linear Precoding Using Channel Estimation for MIMO-OFDM Systems	735
R. Raja Kumar, R. Pandian, C. Satheeswaran, M. KaviPriya and P. Indumathi	
Development of Sine Cosine Toolbox for LabVIEW	747
Shubham Maurya, Mohit Jain and Nikhil Pachauri	
A Survey on Big Data in Healthcare Applications	755
M. Ambigavathi and D. Sridharan	
Solar-Assisted Cloud-Enabled Intelligent Mini Refrigerator	765
Nishi Khetwal, Basant Singh Bhaskar, Udayveer Mittal, Sourav Thakur, Vindhya Devalla, Adesh Kumar and Vivek Kaundal	
Internet of Things in Home Automation—A Review	773
Aayushi, A. K. Nida Susan, Natasha Sanjay Ayare, N. S. Prema and S. B. Chandini	
A Study on Different Types of Web Crawlers	781
P. G. Chaitra, V. Deepthi, K. P. Vidyashree and S. Rajini	

Security and Trust in IoT Communications: Role and Impact	791
Keshav Kaushik and Kamalpreet Singh	
Optimization of Log Gabor Filters Using Genetic Algorithm for Query by Image Content Systems	799
N. Jyothi, D. Madhavi and M. R. Patnaik	
Self-balancing of a Bike Using Gyroscopic Force and PID Controller	807
Animesh Jain, Satyam Bhaskar, Kunal Nandanwar and Hari Om Bansal	
Parallel DES with Modified Mode of Operation	819
Kinjal Chaudhari and Payal Prajapati	
A Novel Compact UWB Antenna for Wireless Applications	833
Vishal Pant, Yogendra Choudhary, Abhinav Sharma, Tarun Kumar and R. Gowri	
Monitoring and Control of City Water Distribution System	841
Shubham Kumar, Nishant Kumar and Arpit Jain	
Detection of DTMF by Using Goertzel Algorithm and Optimized Resource-Sharing Approach	851
B. Khaleelu Rehman, Adesh Kumar, Salauddin Mohammad, Mudasar Basha and K. Venkata Siva Reddy	
Combined Markov Model and Zero Watermarking for Integrity Verification of English Text Documents	857
Fatek Saeed and Anurag Dixit	
Polymer Nanocomposite Membranes for Water Treatment	865
Madhu Sharma and Anupam Sharma	
Analysis and Comparison of ECG Signal Quality Assessments Methods	875
Ramesh Babu Chukka and Ch. Sumanth Kumar	
Mobile Sensors Deployment Methods: A Review	883
P. Thrilochan Sharma, Sushabhan Choudhury and Vinay Chowdary	
DEA Model for Exogenously Fixed Variables Under Fuzzy Environment	895
Shivi Agarwal and Trilok Mathur	
Status of Small Hydropower in Himachal Pradesh: A Review	903
Kamal Kashyap, Robin Thakur, Anil Kumar, Raj Kumar Saini and Amar Raj Singh Suri	

Implementation of HMI Using VISUAL BASIC .NET for Induction Hardening Process	913
Himanshu Sharma, Roushan Kumar, Adesh Kumar, Mukul Kumar Gupta and Vivek Kaundal	
A Method to Reduce the Effect of Inter-Sub-band Interference (ISBI) and Inter Carrier Interference (ICI) in UFMC	921
Rajarao Manda and R. Gowri	
Real-Time Coordination with Electromagnetic Relay and Advanced Numerical Relays in Laboratory Environment/Setup	935
Isaac Ramalla, Ram Mohan Sharma and M. A. Ahmed	
Symmetric Cryptography and Hardware Chip Implementation on FPGA	945
Priyanshi Vishnoi, S. L. Shimi and Adesh Kumar	
Design of Microstrip Antenna with Extended Ground	957
N. Prasanthi kumari, P. Suresh Kumar, G. M. G. Madhuri and B. Praveen Kitti	
A Dual-Band High-Gain Patch Antenna with Reduced Cross-Polarization Levels Designed for X Band and K Band	965
Tanisha Gupta, Tarun Kumar and Rajeev Kamal	
Review Analysis of Single-Electron Transistor	975
Raghav Bansal	
Author Index	985

Editors and Contributors

About the Editors

Dr. Sushabhan Choudhury is currently working as Professor in the Department of Electrical and Electronics Engineering, School of Engineering at University of Petroleum and Energy Studies, Dehradun, Uttarakhand (India). He has around 27+ years of experience in academics and research. Dr. Choudhury has received his Ph.D. (in Electronics Engineering) from University of Petroleum and Energy Studies, Dehradun, M.Tech. (in Electronic Design & Technology) from Tezpur Center University (Gold Medalist) and B.Tech. (in Electronics & Communication Engineering) from NIT, Silchar. Dr. Choudhury has authored 52 international refereed journal papers, 17 international conference papers, 11 Indian patents, 5 books, and 3 book chapters. He is also a life member of ISTE. Dr. Choudhury was the Editor for the first two editions of the conference proceedings published with AISC Book Series, Springer.

Dr. Ranjan Mishra is currently working as an Associate Professor in the Department of Electrical and Electronics Engineering, School of Engineering at University of Petroleum and Energy Studies, Dehradun, Uttarakhand (India). He has around 19+ years of experience in academics and research. Dr. Mishra has received his Ph.D. in Microstrip Antenna Design from University of Petroleum and Energy Studies, Dehradun in the year 2016. Dr. Mishra is an author of 20 international refereed journal and international conference papers and 4 edited books. Dr. Mishra has recently edited two special editions for IJET (UAE) Scopus Indexed Journal.

Prof. Raj Gaurav Mishra is currently working as an Assistant Professor (Selection Grade) in the Department of Electrical and Electronics Engineering at the School of Engineering, University of Petroleum and Energy Studies, Dehradun, Uttarakhand (India). He is having around 12+ years of experience in academics and research. Prof. Mishra has received his M.S. in Space Engineering from the

Department of Space Studies, Umea University, Kiruna, Sweden in the year 2007. He is currently pursuing his Ph.D. thesis in Microstrip Antenna Design at University of Petroleum and Energy Studies, Dehradun. Mr. Mishra is an author of 31 international refereed journal, international conference papers and book chapters, 1 Indian patent and 3 edited books. Mr. Mishra has recently edited two special editions for IJET (UAE).

Dr. Adesh Kumar is currently working as Assistant Professor (Selection Grade) in the Department of Electrical and Electronics Engineering at the School of Engineering, University of Petroleum and Energy Studies, Dehradun, Uttarakhand (India). Adesh Kumar has received his Ph.D. in Electronics Engineering from University of Petroleum and Energy Studies, Dehradun, India, M.Tech. (Hons) in Embedded Systems Technology from SRM University, Chennai, B.Tech. in Electronics and Communication Engineering from UPTU Lucknow (Veera Eng., College, Bijnor). He has also worked as Senior Engineer in TATA ELXSI Limited Bangalore and faculty member in ICFAI University, Dehradun, India. His areas of interest are Wireless communications, VLSI Design, Microprocessors and Embedded System. He has published more than 50 research papers in the international conferences and referred journals. He has contributed in many international conferences in different capacity and conference secretary for ICICCD -2016, ICICCD-2017 in UPES.

Contributors

Aayushi Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka, India

Shivi Agarwal Department of Mathematics, BITS Pilani, Pilani, Rajasthan, India

Alok Aggarwal University of Petroleum and Energy Studies, Dehradun, India

Anjali Agrawal Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Sunil Agrawal UIET, Panjab University, Chandigarh, India

M. A. Ahmed ACE, Aurangabad, India

Neelu Jyoti Ahuja University of Petroleum and Energy Studies, Dehradun, India

M. Ambigavathi Department of Electronics and Communication Engineering, CEG Campus, Anna University, Chennai, India

Karan Anand Department of Aerospace Engineering, University of Petroleum and Energy Studies, Dehradun, India

Oleksandr Anishchenko Pryazovskyi State Technical University, Mariupol, Ukraine

Monika Arora Department of IT/Operations, Apeejay School of Management, Dwarka, New Delhi, India

Shubham Ashish Electrical and Electronics Department, KIET, Ghaziabad, India

Neha Awasthi UIET, Panjab University, Chandigarh, India

Natasha Sanjay Ayare Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka, India

Ayush Department of Electrical Engineering, Incubation Center for Control Dynamical Systems and Computation, Delhi Technological University, New Delhi, Delhi, India

Mudit Babel Department of Electronics and Telecommunication Engineering, Rajiv Gandhi College of Engineering and Research, Nagpur, India

Abhishek Baghel Computer Science Department, Bharati Vidyapeeth's College of Engineering, New Delhi, India

Elena Balalayeva Pryazovskyi State Technical University, Mariupol, Ukraine

Hari Om Bansal Department of Electrical and Electronics Engineering, Birla Institute of Technology and Science, Pilani, Rajasthan, India

Kritika Bansal National Institute of Technology Delhi, New Delhi, India

Raghav Bansal Department of Electrical and Electronics, BITS Pilani, Pilani Campus, Pilani, Rajasthan, India

Mudasar Basha BVRIT, Hyderabad, Telangana, India

Harshad Bhadka Faculty of Computer Science, C U Shah University, Wadhwan, Gujarat, India

Sanjay Bhardwaj Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi, South Korea

Shanu Bhardwaj Chitkara University Institute of Engineering and Technology, Chitkara University, Chandigarh, India

Vishal Bharti Department of Computer Science and Engineering, DIT University, Dehradun, Uttarakhand, India

Basant Singh Bhaskar Department of Mechanical Engineering, University of Petroleum and Energy Studies, Dehradun, India

Satyam Bhaskar Department of Electrical and Electronics Engineering (B. E. (Hons.)), Department of Physics (M. Sc. (Hons.)), Birla Institute of Technology and Science, Pilani, India

Amandeep Bisht UIET, Panjab University, Chandigarh, India

Mangal Singh Bisht Department of Basic Sciences, GBPE College, New Delhi, Uttarakhand, India

Saji Chacko Electrical Power System, SSTC, SSGI's, Bhilai, India; Electrical Engineering, Government Polytechnic, Durg, Chhattisgarh, India

P. G. Chaitra Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka, India

Prasun Chakrabarti Department of Computer Science Engineering, ITM Universe, Vadodara, Gujarat, India

Vivek Chamoli Graphic Era (Deemed to Be University), Dehradun, India

S. B. Chandini Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka, India

Kinjal Chaudhari Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Piyush Chauhan Department of Analytics, School of Computer Science, University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India

Akanksha Bansal Chopra Department of Computer Science, SPM College, University of Delhi, New Delhi, India

Yogendra Choudhary Department of Electronics Instrumentation & Control Engineering, University of Petroleum and Energy Studies, Dehradun, India

Sushabhan Choudhury Electrical and Electronics Department, University of Petroleum and Energy Studies, Dehradun, India

Vinay Chowdary Electrical and Electronics Department, University of Petroleum and Energy Studies, Dehradun, India

Ramesh Babu Chukka Department of Electronics and Communication Engineering, Vignan's Institute of Engineering for Women, Visakhapatnam, India

Ratna Dahiya Department of Electrical Engineering, National Institute of Technology, Kurukshetra, India

Arindam Deb Department of Computer Science and Engineering, National Institute of Technology, Agartala, Barjala, Jirania, India

Praseed Debbarma Department of Computer Science and Engineering, National Institute of Technology, Agartala, Barjala, Jirania, India

V. Deepthi Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka, India

Vindhya Devalla Department of Aerospace Engineering, University of Petroleum and Energy Studies, Dehradun, India

Mishra Devesh Department of Electronics & Communication, University of Allahabad, Allahabad, India

Abhiraj Dharavath Department of Electrical Engineering, Incubation Center for Control Dynamical Systems and Computation, Delhi Technological University, New Delhi, Delhi, India

Anaahat Dhindsa University Institute of Engineering and Technology, Panjab University, Chandigarh, India

Anurag Dixit Computer Science and Engineering, Galgotias University, New Delhi, India

Veer Sain Dixit Department of Computer Science, ARSD College, University of Delhi, New Delhi, India

Daksh Dobhal Department of Electrical Engineering, Incubation Center for Control Dynamical Systems and Computation, Delhi Technological University, New Delhi, Delhi, India

Ayush Dogra UIET, Panjab University, Chandigarh, India

Sagnik Dutta Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies, Dehradun, India

Subhankar Dutta Department of Electrical and Electronics Engineering, SRM Institute of Science and Technology, Chennai, India

R. P. Dwivedi Shoolini University, Solan, Himachal Pradesh, India

Niranjan Gandigudi Department of Aerospace Engineering, University of Petroleum and Energy Studies, Dehradun, India

Neeraj Gандотра Shoolini University, Solan, Himachal Pradesh, India

Sagar Garg Electrical and Electronics Department, KIET, Ghaziabad, India

Pratik Ghosh Murshidabad College of Engineering and Technology, Murshidabad, WB, India

Arnav Goel Department of Electrical Engineering, Incubation Center for Control Dynamical Systems and Computation, Delhi Technological University, New Delhi, Delhi, India

Naveen Goel Department of Electrical and Electronics Engineering, SSTC, Bhilai, Chhattisgarh, India

Somya Goel Computer Science Department, Bharati Vidyapeeth's College of Engineering, New Delhi, India

Veerapu Goutham Department of Electronics and Communication Engineering, National Institute of Technology Puducherry, Karaikal, India

R. Gowri Department of Electronics Instrumentation & Control Engineering, University of Petroleum and Energy Studies, Dehradun, India

Bhawna Goyal UIET, Panjab University, Chandigarh, India

Piyush Goyal Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India

Mukul Kumar Gupta Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies (UPES), Dehradun, Uttarakhand, India

Rajeev Gupta University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India

Shweta Gupta Electronics and Communication Engineering Department, Jain University, Bangalore, India

Tanisha Gupta Department of Electronics and Electrical Engineering, University of Petroleum and Energy Studies, Dehradun, India

S. L. Haridas Department of Electronics & Telecommunication Engineering, G. H. Raisoni College of Engineering, Nagpur, India

V. P. Harigovindan Department of Electronics and Communication Engineering, National Institute of Technology Puducherry, Karaikal, India

S. A. Hariprasad Electronics and Communication Engineering Department, Jain University, Bangalore, India

Harshad Harsoda Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Naimul Hasan Sanaka Educational Trust's Group of Institution, Durgapur, WB, India

Akshay Hinduja Department of Computer Applications, National Institute of Technology Raipur, Raipur, Chhattisgarh, India

Svitlana Hurkovska Donbas State Engineering Academy, Kramatorsk, Ukraine

P. Indumathi Department of Electronics Engineering, Anna University, MIT Campus, Chennai, India

M. Jagabar Sathik Department of Electrical and Electronics Engineering, SRM Institute of Science and Technology, Chennai, India

Animesh Jain Department of Mechanical Engineering (B. E. (Hons.)), Department of Biological Science (M. Sc. (Hons.)), Birla Institute of Technology and Science, Pilani, India

Anuj Jain Computer Science Dept, Institute of Technology and Management Universe, Vadodara, Gujarat, India

Arpit Jain UPES, Dehradun, Uttarakhand, India

Mohit Jain School of Automation, Banasthali Vidyapith, Vanasthali, Rajasthan, India

R. Janani Department of Electronics and Instrumentation Engineering, Sri Chandrasekharendra Sarawathi Viswa Maha Vidyalaya, SCSVMV, Enathur, Kanchipuram, TamilNadu, India

Vibhu Jately Department of ECE, KIoT, Wollo University, Kombolcha, Ethiopia

S. Jayamangala ECE Department, J.B. Institute of Engineering and Technology, Hyderabad, India

Upasna Joshi Department of Computer Science and Engineering, DTC, Greater Noida, India

N. Jyothi Department of EIE, GIT, GITAM Deemed to be University, Visakhapatnam, Andhra Pradesh, India

Rajeev Kamal Department of Electronics and Communication Engineering, Dayananda Sagar University, Bengaluru, Karnataka, India

Agrawal Krishna Kant Department of Computer Science & Engineering, Delhi Technical Campus, Greater Noida, India

Kumar Kartikeya Department of Electrical Engineering, Incubation Center for Control Dynamical Systems and Computation, Delhi Technological University, New Delhi, Delhi, India

Kamal Kashyap Department of Mechanical Engineering, A P Goyal Shimla University, Shimla, HP, India

K. Kathiravan Department of Information Technology, Easwari Engineering College, Chennai, Tamilnadu, India

Vivek Kaundal Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies, Dehradun, India

Amanpreet Kaur Antenna Research and Testing Laboratory, Department of Electronics and Communication Engineering, Thapar University, Patiala, India; ECED, TIET Patiala, Punjab, India

Arashpreet Kaur Antenna Research and Testing Laboratory, Department of Electronics and Communication Engineering, Thapar University, Patiala, India

Chamandeep Kaur UIET, Panjab University, Chandigarh, India

Gagandeep Kaur ECED, TIET Patiala, Punjab, India

Manpreet Kaur Department of Electrical Engineering, Incubation Center for Control Dynamical Systems and Computation, Delhi Technological University, New Delhi, Delhi, India

Parminder Kaur Department of Computer Science, Guru Nanak Dev University, Amritsar, India

Keshav Kaushik School of Computer Science, UPES, Dehradun, India

Shiva Kaushik Electronics and Communication Engineering Department, G.L.A University, Mathura, U.P, India

M. KaviPriya Department of Electronics Engineering, Anna University, MIT Campus, Chennai, India

Pragati Kene Department of Electronics Engineering, G.H. Raisoni College of Engineering, Nagpur, India

Kishore Keshav Department of Computer Science and Engineering, Alakh Prakash Goyal Shimla University, Shimla, Himachal Pradesh, India

Vipin Khattri Department of Computer Application, Integral University, Lucknow, India

Nishi Khetwal Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies, Dehradun, India

Dong-Seong Kim IEEE, Gumi, South Korea

Nakshatra Kothi Computer Science Dept, Institute of Technology and Management Universe, Vadodara, Gujarat, India

Alok Darshan Kothiyal Department of Mathematics, Doon College Rishikesh, Rishikesh, Uttarakhand, India

Nakul Kotibhaskar New Delhi, India

Volodymyr Kukhar Pryazovskyi State Technical University, Mariupol, Ukraine

Adarsh Kumar University of Petroleum and Energy Studies, Dehradun, India

Adesh Kumar Department of Electrical and Electronics Engineering, School of Engineering, University of Petroleum and Energy Studies, Dehradun, India

Amit Kumar School of Computer Science and Engineering, University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India

Anil Kumar School of Mechanical & Civil Engineering, Shoolini University, Solan, HP, India

Ch. Sumanth Kumar Department of Electronics and Communication Engineering, GITAM Institute of Technology, GITAM Deemed to Be University, Visakhapatnam, India

D. Kumar University of Petroleum & Energy Studies, Dehradun, India

Megha Kumar Department of Computer Science and Engineering, DTC, Greater Noida, India

Nishant Kumar UPES, Dehradun, Uttarakhand, India

Ravinder Kumar Department of Electrical and Electronics Engineering, Birla Institute of Technology and Science, Pilani, Rajasthan, India

Ravindra Kumar Department of Computer Science and Engineering, DTC, Greater Noida, India

Roushan Kumar Department of Mechanical Engineering, School of Engineering, University of Petroleum and Energy Studies Dehradun, Dehradun, India

Sanjeev Kumar UIET, Panjab University, Chandigarh, India

Shashwat Kumar Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies, Dehradun, India

Shubham Kumar UPES, Dehradun, Uttarakhand, India

Shukla Narendra Kumar Department of Electronics & Communication, University of Allahabad, Allahabad, India

Sunil Kumar Department of Mathematics, UTU, Dehradun, India;
School of Mechanical & Civil Engineering, Shoolini University, Solan, India

Tarun Kumar Department of Electronics Instrumentation & Control Engineering, University of Petroleum and Energy Studies, Dehradun, India

Vineet Kumar Department of Electrical Engineering, Kurukshetra University, Kurukshetra, India

Ch. Aruna Kumari Department of ECE, PSCMR College of Engineering, Vijayawada, Andhra Pradesh, India

N. Prasanthi Kumari University of Petroleum and Energy Studies, Dehradun, India

V. V. Kuppast Mechanical Engineering Department, Basveshwar College of Engineering, Bagalkot, Karnataka, India

S. N. Kurbet Mechanical Engineering Department, Basveshwar College of Engineering, Bagalkot, Karnataka, India

Caneon Kurien Department of Mechanical Engineering, University of Petroleum and Energy Studies, Dehradun, India

Pradeep Laxkar Department of Computer Science Engineering, ITM Universe, Vadodara, Gujarat, India

D. Madhavi Department of EIE, GIT, GITAM Deemed to be University, Visakhapatnam, Andhra Pradesh, India

Charu Madhu University Institute of Engineering and Technology, Panjab University, Chandigarh, India

G. M. G. Madhuri Department of ECE, PSCMR College of Engineering, Vijayawada, Andra Pradesh, India

Ravi Prakash Mahobia Department of Electrical and Electronics Engineering, SSTC, Bhilai, Chhattisgarh, India

Monu Malik Department of Electronics and Electrical Engineering, Maharaja Surajmal Institute of Technology, New Delhi, India

Rajaraao Manda Department of Electrical and Electronics Engineering, UPES, Dehradun, India

Manuja Manoj Vertical Head (IT) Inurture Education Solutions, Bangalore, India

Oleg Markov Donbas State Engineering Academy, Kramatorsk, Ukraine

Anu Marwaha Chitkara University Institute of Engineering and Technology, Chitkara University, Chandigarh, India

Trilok Mathur Department of Mathematics, BITS Pilani, Pilani, Rajasthan, India

Vishal Mathur The ICFAI University, Jaipur, India

Shubham Maurya Department of Electronics and Communication, GLA University, Mathura, India

C. S. Meera Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies (UPES), Dehradun, Uttarakhand, India

Jagriti Mishra Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies (UPES), Dehradun, India

Urvi Mitra Amity University, Lucknow, India

Udayveer Mittal Department of Mechanical Engineering, University of Petroleum and Energy Studies, Dehradun, India

Salauddin Mohammad ECE Department, J.B. Institute of Engineering and Technology, Hyderabad, Telangana, India

Amit Kumar Mondal Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies, Dehradun, India

Ch. Monica Research Scholar KL University, Vaddeswaram, India

Manishita Mukherji Amity University Jaipur, Jaipur, Rajasthan, India

Rana Mukherji The ICFAI University, Jaipur, India

Pankaj Mukhija National Institute of Technology Delhi, New Delhi, India

Dhara Mungra Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Atanu Nag Modern Institute of Engineering & Technology, Bandel, Hooghly, WB, India

Vonteddu Shanmukha NagaRaju Department of Electrical and Electronics Engineering, Jawaharlal Nehru Technological University, Kakinada, India

Preeti Nagrath Computer Science Department, Bharati Vidyapeeth's College of Engineering, New Delhi, India

Kunal Nandanwar Department of Mechanical Engineering (B. E. (Hons)), Birla Institute of Technology and Science, Pilani, India

Sandeep Kumar Nayak Department of Computer Application, Integral University, Lucknow, India

Aradhana Negi Department of Computer Science, Guru Nanak Dev University, Amritsar, India

A. K. Nida Susan Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka, India

S. Nithyanandam Department of Computer Science and Engineering, Prist University, Vallam, Thanjavur, India

Nikhil Pachauri Department of Electrical and Instrumentation, Thapar Institute of Engineering and Technology, Patiala, India

Siddharth Pachpor Department of Electrical and Electronics Engineering, SRM Institute of Science and Technology, Chennai, India

Tannistha Pal Department of Computer Science and Engineering, National Institute of Technology, Agartala, Barjala, Jirania, India

S. N. Panda Chitkara University Institute of Engineering and Technology, Chitkara University, Chandigarh, India

Alok Kumar Pandey Electrical and Electronics Department, KIET, Ghaziabad, India

Jitender Pandey Graphic Era (Deemed to Be University), Dehradun, India

Manju Pandey Department of Computer Applications, National Institute of Technology Raipur, Raipur, Chhattisgarh, India

Rahul Pandey Electrical Power System, SSTC, SSGI's, Bhilai, India

Surya Pandey Orchids The International School, Sarjapur, Bangalore, Karnataka, India

R. Pandian Department of Electronics and Instrumentation Engineering, Sathyabama Institute of Science and Technology, Chennai, India

Vishal Pant Department of Electronics Instrumentation & Control Engineering, University of Petroleum and Energy Studies, Dehradun, India

Shailesh Singh Panwar H.N.B. Garhwal University, Srinagar, Garhwal, Uttarakhand, India

M. R. Patnaik Department of Instrument Technology, AU College of Engineering, Andhra University, Visakhapatnam, Andhra Pradesh, India

Payal Prajapati Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Rishi Prakash Graphic Era (Deemed to Be University), Dehradun, India

N. Prasanthi kumari Department of Electronics, Department of Mechanical, University of Petroleum and Energy Studies, Dehradun, India

B. Praveen Kitti Department of ECE, PSCMR College of Engineering, Vijayawada, Andhra Pradesh, India

N. S. Prema Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka, India

Kocherlakota Satya Pritam Department of Mathematics, BITS Pilani, Pilani, Rajasthan, India

Rashmi Priya Asansol Engineering College, Asansol, WB, India

Andrii Prysiazhnyi Pryazovskyi State Technical University, Mariupol, Ukraine

Sonal Purkait Electrical Power System, SSTC, SSGI's, Bhilai, India

R. Radha Department of Computer Science and Engineering, K L University, Hyderabad, Telangana, India

Y. P. Raiwani H.N.B. Garhwal University, Srinagar, Garhwal, Uttarakhand, India

R. Raja Kumar Mathematics Department, Sathyabama Institute of Science and Technology, Chennai, India

S. Rajini Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka, India

Isaac Ramalla MLRITM, Hyderabad, India

Muhammad Rusyadi Ramli Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi, South Korea

Rayapudi Srinivasa Rao Department of Electrical and Electronics Engineering, Jawaharlal Nehru Technological University, Kakinada, India

Gunjan Rathore University Institute of Engineering and Technology, Panjab University, Chandigarh, India

Manam Ravindra Department of Electrical and Electronics Engineering, Jawaharlal Nehru Technological University, Kakinada, India

S. R. N. Reddy Department of Computer Science and Engineering, Indira Gandhi Delhi Technical University for Women, Delhi, India

B. Khaleelu Rehman Department of Electrical & Electronins Engineering, Univeristy of Petroleum & Energy Studies (UPES), Dehradun, Uttarakhand, India

G. Rekha Department of Computer Science and Engineering, K L University, Hyderabad, India

A. Rijuvana Begum Research Guide, Department of E.C.E, Prist University, Thanjavur, India

Anwesha Roy Department of Computer Science and Engineering, National Institute of Technology, Agartala, Barjala, Jirania, India

Kousik Roy Asansol Engineering College, Asansol, WB, India

Fatek Saeed Computer Science and Engineering, Galgotias University, New Delhi, India

Yuri Sahirov Pryazovskyi State Technical University, Mariupol, Ukraine

Lalit kumar Sahu Electrical Power System, SSTC, SSGI's, Bhilai, India

Devender Kumar Saini University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India

Raj Kumar Saini School of Mechanical and Civil Engineering, Shoolini University, Solan, Himachal Pradesh, India

Amandeep Singh Sappal Punjabi University, Patiala, India

Sagar Saraswat Electrical and Electronics Department, KIET, Ghaziabad, India

Prasanta Sarkar Department of Electrical Engineering, National Institute of Technical Teachers' Training and Research, Kolkata, West Bengal, India

C. Satheswaran Department of Electronics and Communication Engineering, Dhaanish Ahmed College of Engineering, Chennai, India

R. Seetha ECE Department, J.B. Institute of Engineering and Technology, Hyderabad, India

Dikshan N. Shah Gujarat Technological University, Ahmedabad, Gujarat, India

Kunal Shah Department of Computer Science Engineering, ITM Universe, Vadodara, Gujarat, India

Shashvat Shah Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Sharma Shamneesh Department of Computer Science and Engineering, Chitkara University, Chandigarh, Punjab, India

Zeenat Shareef Department of Electronics and Communication Engineering, Indira Gandhi Delhi Technical University for Women, Delhi, India

Abhinav Sharma Department of Electronics Instrumentation & Control Engineering, University of Petroleum and Energy Studies, Dehradun, India

Anupam Sharma Center for Energy Studies, Indian Institute of Technology Delhi, Hauz Khas, New Delhi, India

Apoorav Maulik Sharma UIET, Panjab University, Chandigarh, India

Ashwani Sharma Shoolini University, Solan, Himachal Pradesh, India

Himanshu Sharma Department of Electrical and Electronics, School of Engineering, University of Petroleum and Energy Studies, Dehradun, India

Madhu Sharma Department of Electrical Engineering, University of Petroleum and Energy Studies, Dehradun, India

Ram Mohan Sharma UPES, Dehradun, India

Ritu Sharma Department of Electrical and Electronics Engineering, SSTC, Bhilai, Chhattisgarh, India

Shalini Sharma School of Electrical and Computer Sciences, Shoolini University, Solan, Himachal Pradesh, India

Sukesha Sharma UIET, Panjab University, Chandigarh, India

Jeel Sheth Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

S. L. Shimi Department of Electrical Engineering, National Institute of Technical Teachers Training and Research (NITTTR), Chandigarh, India

Aasheesh Shukla Electronics and Communication Engineering Department, G.L. A University, Mathura, U.P, India;

Department of Electronics and Communication, GLA University, Mathura, India

Deepak Kumar Singh Jaipuria Institute of Management, Lucknow, India

Dheerendra Singh Department of Electrical and Electronics Engineering, Birla Institute of Technology and Science, Pilani, Rajasthan, India

Harmeet Singh Punjabi University, Patiala, India;
Chandigarh University, Mohali, India

Inder Singh School of Computer Science Engineering, UPES, Dehradun, UK

Juhি Singh Department of Electronics and Communication, GLA University, Mathura, India

Kamalpreet Singh School of Computer Science, UPES, Dehradun, India

Krishna Veer Singh Department of Electrical and Electronics Engineering, Birla Institute of Technology and Science, Pilani, Rajasthan, India

Lairenlakpam Joyprakash Singh Department of Electronics & Communication Engineering, School of Technology, North-Eastern Hill University, Shillong, Meghalaya, India

Madhusudan Singh Department of Electrical Engineering, Incubation Center for Control Dynamical Systems and Computation, Delhi Technological University, New Delhi, Delhi, India

Preeti Singh UIET, Panjab University, Chandigarh, India

Shikha Singh Department of Computer Science and Engineering, Amity School of Engineering and Technology, Amity University, Lucknow, Uttar Pradesh, India

Ravi Singhal University of Petroleum and Energy Studies, Dehradun, India

Anshu Singla Chitkara University Institute of Engineering and Technology, Chitkara University, Chandigarh, India

Asefa Sisay Department of ECE, KIoT, Wollo University, Kombolcha, Ethiopia

N. Sreenath Department of Computer Science and Engineering, Pondicherry Engineering College, Puducherry, India

D. Sridharan Department of Electronics and Communication Engineering, CEG Campus, Anna University, Chennai, India

Ajay Kumar Srivastava Department of Mechanical Engineering, University of Petroleum and Energy Studies, Dehradun, India

Aprajita Srivastava Computer Science Department, Bharati Vidyapeeth's College of Engineering, New Delhi, India

Garima Srivastava Amity University, Lucknow, India

Tanya Srivastava Department of Computer Science and Engineering, Amity School of Engineering and Technology, Amity University, Lucknow, Uttar Pradesh, India

Yadav Ram Suchit Department of Electronics & Communication, University of Allahabad, Allahabad, India

P. Suresh Kumar Department of Electronics, Department of Mechanical, University of Petroleum and Energy Studies, Dehradun, India

Amar Raj Singh Suri School of Mechanical and Civil Engineering, Shoolini University, Solan, Himachal Pradesh, India

Jaydeep Swarnakar Department of Electronics & Communication Engineering, School of Technology, North-Eastern Hill University, Shillong, Meghalaya, India

Basavaraj Talikoti Mechanical Engineering Department, Dr. K. M. Vasudevan Pillai College of Engineering Media Studies and Research, New Panvel, Maharashtra, India

Pande Tanuja Department of Electronics & Communication, University of Allahabad, Allahabad, India

Ankit Thakkar Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Robin Thakur School of Mechanical & Civil Engineering, Shoolini University, Solan, Himachal Pradesh, India

Sourav Thakur Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies, Dehradun, India

Sashank Thapa School of Mechanical & Civil Engineering, Shoolini University, Solan, HP, India

I. Thirunavukkarasu Department of Instrumentation and Control Engineering, Manipal Academy of Higher Education, Manipal, Karnataka, India

P. Thrilochan Sharma Automation and Robotics, University of Petroleum and Energy Studies, Dehradun, India

S. M. Towhidul Islam Nayim Department of Mechanical Engineering, A P Goyal Shimla University, Shimla, HP, India

Aayushi Tyagi Computer Science Department, Bharati Vidyapeeth's College of Engineering, New Delhi, India

Amit Kumar Tyagi School of Computer Science and Engineering, Lingayas Vidyapeeth, Faridabad, Haryana, India

Pradeep Udupa Research Scholar, Department of C.S.E, Prist University, Thanjavur, India

Shivansh Upadhyay Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies, Dehradun, India

Sudarshan K. Valluru Department of Electrical Engineering, Incubation Center for Control Dynamical Systems and Computation, Delhi Technological University, New Delhi, Delhi, India

Vineeth Veeramalla Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies (UPES), Dehradun, India

K. Venkata Siva Reddy Ravindra College of Engineering for Women, Kurnool, Andhra Pradesh, India

Anjani Kumar Verma Department of Computer Science, University of Delhi, Delhi, India

Piush Verma Rayat-Bahra Group of Institutes, Chandigarh, Punjab, India

Varnita Verma Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies (UPES), Dehradun, India

K. P. Vidyashree Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, Karnataka, India

Renu Vig UIET, Panjab University, Chandigarh, India

Priyanshi Vishnoi Department of Electrical Engineering, National Institute of Technical Teachers Training and Research (NITTTR), Chandigarh, India

Kevin Vora Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Performance Improvement of Three-Phase Squirrel Cage Induction Motor Operating Under Rated Voltages—A Design Consideration for Rural Areas



Raj Kumar Saini, Devender Kumar Saini, Rajeev Gupta, Piush Verma,
Neeraj Gандотра, Ashwani Sharma, Robin Thakur and R. P. Dwivedi

Abstract Because of reactive power requested by the industrialists and agriculturists, three-phase induction motors composed at evaluated voltage are working under the appraised voltage particularly in the rural areas, a long way from the utility focuses. The target of this research article is to enhance torque, effectiveness, power developed, and power factor by using double-cage winding with a new combination of stator and rotor design. The rim_003 rotor bar is used with so_012 parallel tooth opening flat bottom stator with different arrangements of stator winding, which are used for 5 HP three-phase squirrel cage induction motor. JMAG Express is used to figure the execution enhancement of motor under appraised voltage.

Keywords JMAG Express · IM · 5 H.P · Losses · Efficiency · Torque

1 Introduction

Energy protection is an essential development venture toward beating the creating issues of businesses enhancement. By decreasing the wasteful imperativeness, the execution of the electrical machines can be upgraded up to some extent. In Refs. [1, 2], the creators portrayed that the three-phase induction motor accepts a basic occupation for the progression of thing in the enterprises. In agribusiness country like India, a tremendous number of induction motor pump sets are working under the appraised voltage. The capability of these motors can be extended by diminishing the misfortunes, improving the cross-areas of the conductors, and changing the

R. K. Saini (✉) · N. Gандотра · A. Sharma · R. Thakur · R. P. Dwivedi
Shoolini University, Solan 17329, Himachal Pradesh, India
e-mail: rajsaini.acet@gmail.com

R. K. Saini · D. K. Saini · R. Gupta
University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India

P. Verma
Rayat-Bahra Group of Institutes, Chandigarh, Punjab, India

blueprint measurements of the motors according to the assortment of the voltages. In Refs. [3, 4], the makers figured the delayed consequences of three-phase induction with respect to unequal voltages and fixed that in light of this type voltages, both adequacy and lifetime of the motor lessen. In Ref. [5], the creators did 12 variable components which impacted the proficiency of the motor while working under low voltages. In Ref. [6], P. Pillay and others worked on sequences of voltages. In Ref. [7], the creators said that the expanded scope of unequal voltages affected rotor losses when contrasted with stator losses due to more current. In Ref. [8], tests on IM are carried out by considering complex voltage factor. The creator of Ref. [9] explored that the efficiency of the motor can also be expended by expanding the stack length. In Ref. [10], the authors studied the effects of lopsided voltages on three-phase IM. The creator of the Ref. [11] displayed another method on 1.5 kW IM to compute the proficiency and other trademarks by utilizing a bacterial scrounging calculation (BFA). In Ref. [12], the creator exhibited a strategy to quantify the forecast of temperature ascend in the winding of the machine because of harmonics. In Refs. [13, 14], the writers examined in detail the distinctive reasons, focus of recommendations, and meanings of uneven voltages. The authors of Refs. [15–20] investigated the execution and structure alteration in IM by using JMAG Express.

1.1 Impacts of Under Voltages

It is found that according to the perceptions talked about in Table 2, the normal voltage of the three phases is low, which affects the execution of the motors. Basically, there are two sorts of misfortunes that occur in the IM: one is constant misfortune which depends on the development features of the motor and varies in the extent of 2–3%, and another is the variable misfortune which depends on the load and voltage conditions. The commitment of friction and windage losses is just 0.5–1%.

1.2 Three-Phase Induction Motor and JMAG Express

It is an arrangement of coordinated programming software which has an aim to help electric motor originators all through the total plan process. Concept design properties are computed by the fast mode activity in seconds by choosing the stator and rotor design. The required geometry can be gotten effortlessly in the wake of adjusting the parameters as appeared in Fig. 1. A 5 HP three-phase squirrel cage IM has been proposed with the accompanying details as shown in Table 1.

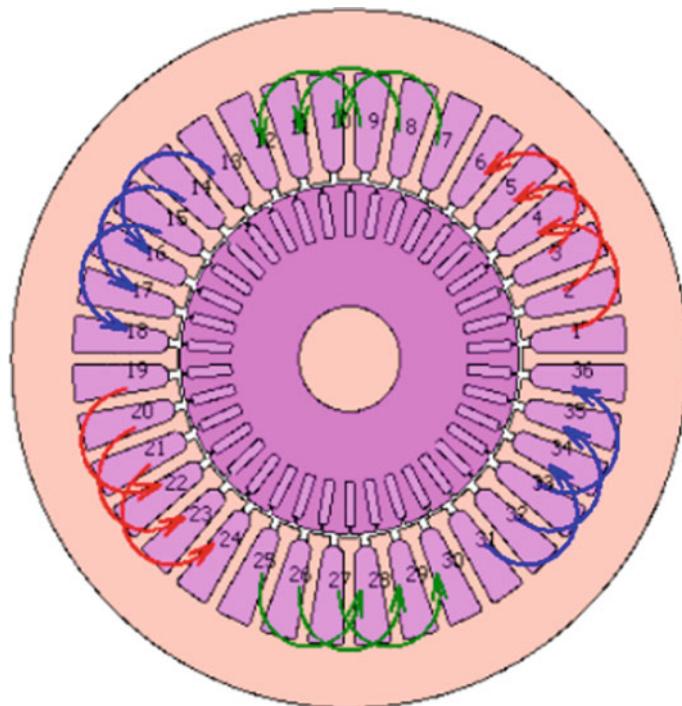


Fig. 1 Stator and rotor design with single-layer winding

Table 1 Information particulars

Particulars	Values
Rating of the machine	5 HP
Rated speed	1440 R.P.M
Maximum speed	1500 R.P.M
Number of poles	4
Number of slots	36
Number of bars	38
Power supply (RMS)	370 V
Maximum current (R.M.S)	10 Amp
Stator design	rim_012 parallel tooth opening, flat bottom
Rotor design	so_003 rectangular bar
Cage	Aluminum
Winding layer	Single/double layer

1.3 Simulation Results and Discussion

The proposed JMAG Express method is used for 5HP three-phase squirrel cage induction motor to figure out the stator and rotor dimensions at maximum torque, efficiency, power developed, and power factor, when the machine works under the evaluated voltage. Working execution of the motor much relies on the outline of stator and rotor design. To decrease the magnetic current rim_012 parallel tooth opening, flat base stator with so_003 rectangular rotor bar is used according to the geometry as appeared in Fig. 1.

Two sorts of windings are used to check the performance of the motor under the rated voltage. In single-layer winding, the space is secured by just a single-loop side. For the most part, this sort of winding cannot be utilized in commutator compose motors, while in the two-layer winding, one curl side of each loop is set in top portion of the space and another loop side is set in lower half of the opening. Torque and power can be enhanced by using two-layer wingdings for the most part when the motors are working under the evaluated voltages. Figures 1 and 4 depict stator and rotor design by using single- and double-layer winding.

Figures 2 and 3 show the stator winding and stator slot winding arrangement, when the winding is connected in star with parallel number 1 and series number 12. Generally, this kind of winding is done in small rating induction motors working under rated voltages. When motors having this sort of winding operate under the rated voltages specific in the villages, the torque and power cannot be procured as

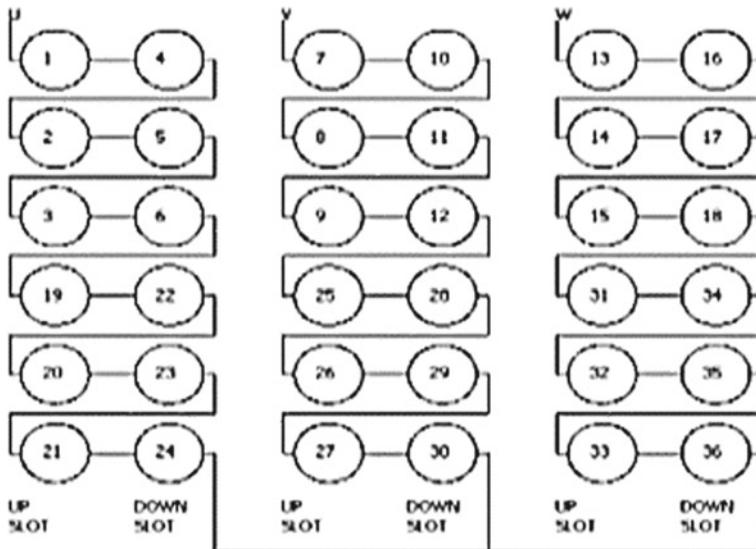


Fig. 2 Single-layer stator winding arrangement with parallel no. 1 and series no. 12 (star connected)

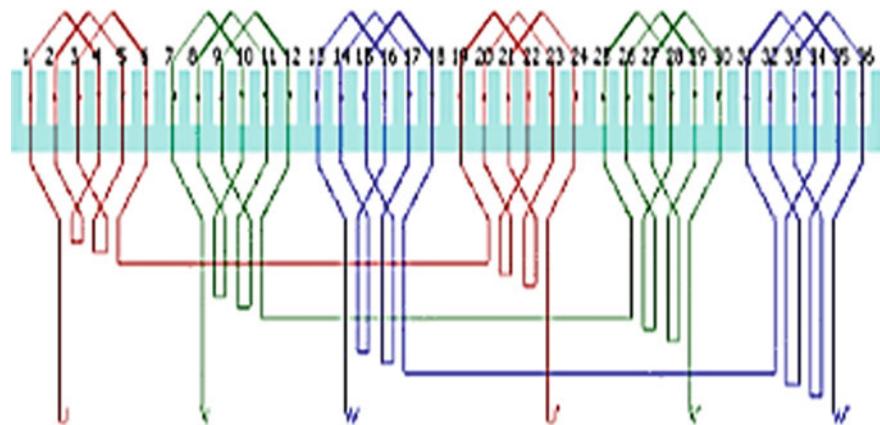


Fig. 3 Single-layer stator slot winding diagram with parallel no. 1 and series no. 12 (star connected)

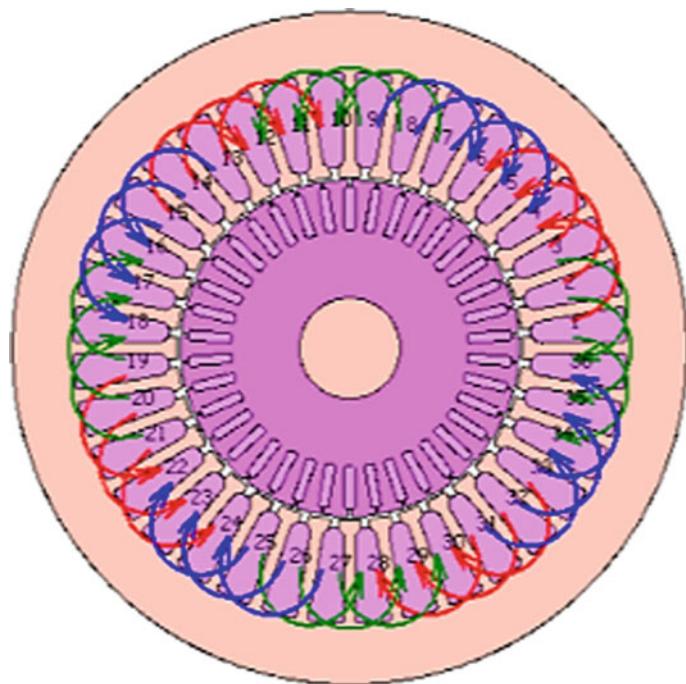


Fig. 4 Double-layer stator and rotor design winding diagram (star connected)

given on the nameplate of the machine, which specifically influence the performance of the motors.

Figure 5 demonstrates the stator slot winding arrangement, when the winding is associated with parallel number 1 and series number 12, while Fig. 6 shows the stator slot winding arrangement for parallel number 2 and series number 6.

Figure 7 demonstrates the torque and speed relationship for single-layer and double-layer winding in two modes. In the first mode, the stator winding is done by making one parallel path and twelve series paths. Under this arrangement of winding when the motor is operating under rated voltage, the starting and maximum torques were 0.430 and 1.187 Nm for single-layer winding while those for double-layer winding were 0.538 and 1.339 Nm. In the second mode, the winding is connected in parallel paths 2 and series paths 6. In this mode, the starting and maximum torques were measured as 0.883 and 2.34 Nm for single-layer winding, but when the double-layer winding is used in this arrangement, the starting and maximum torques were improved up to 1.076 and 2.67 Nm, respectively.

Figure 8 demonstrates the relation between power developed and speed, similarly by the same modes as discussed above; in case of the first mode, the maximum power developed was 195.082 and 175.359 W, but in the second mode, it was raised up to 388.373 and 350.704 W, respectively; however, the starting power developed in case of double-layer winding was much better in both the arrangements.

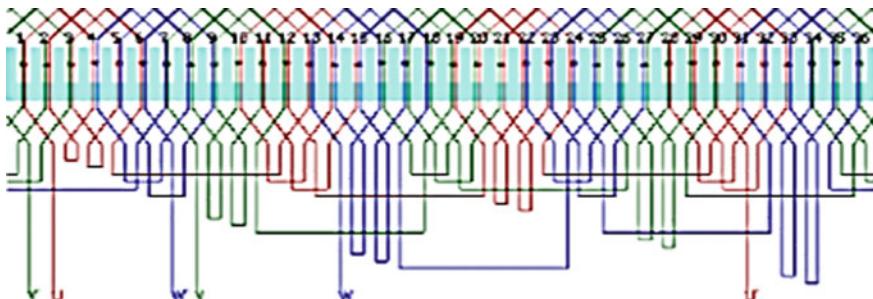


Fig. 5 Double-layer stator slot winding design with parallel no. 1 and series no. 12 (star connected)

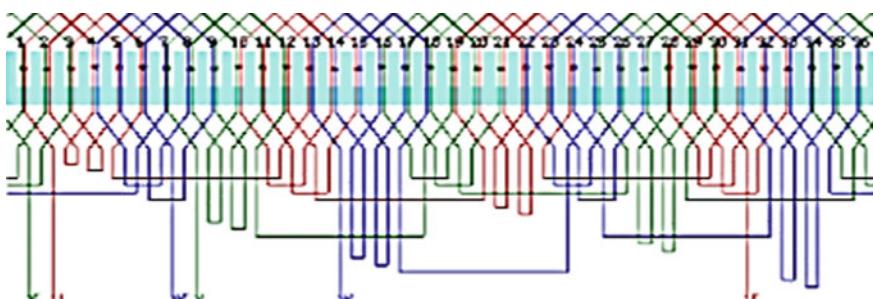


Fig. 6 Double-layer stator slot winding design with parallel no. 2 and series no. 6 (star connected)

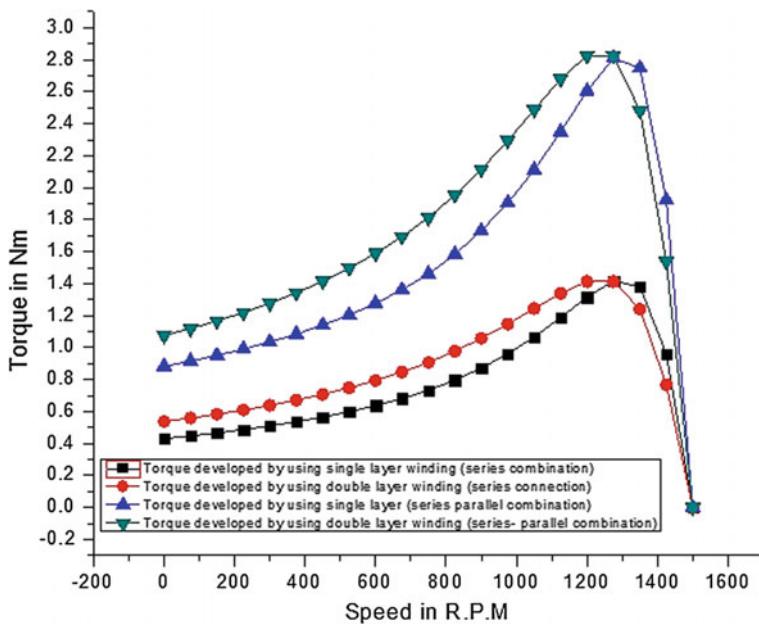


Fig. 7 Torque versus speed

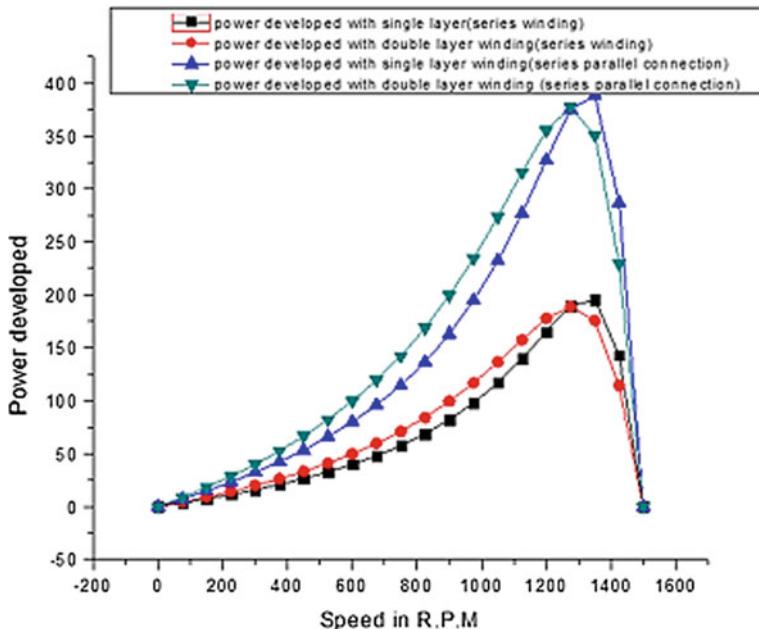


Fig. 8 Power developed versus speed

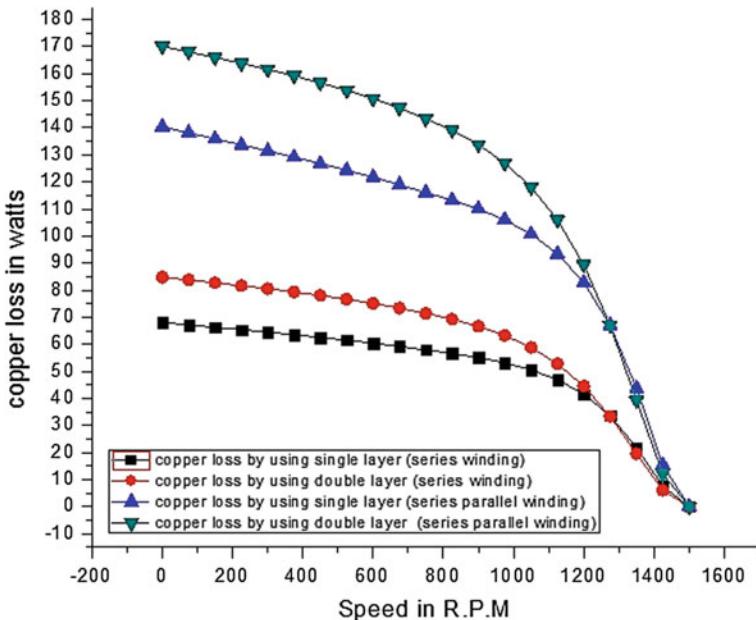


Fig. 9 Copper losses versus speed

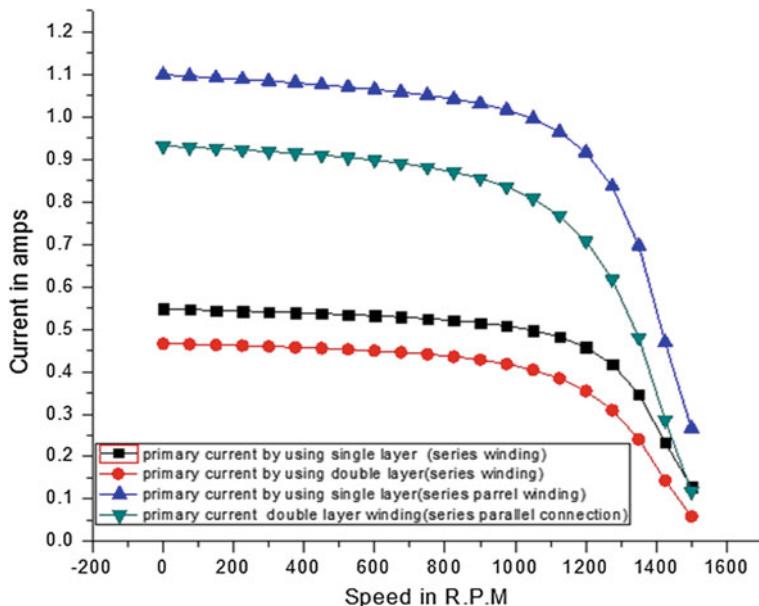
Figure 9 demonstrates copper loss versus speed, similarly by the same mode as discussed above; due to high starting torque and power developed in both the cases, the copper losses will be high at the time of starting of the motor in case of two-layer winding when contrasted with single-layer winding, but at rated speed almost copper losses are the same in both the cases.

While Fig. 10 illustrates the relationship between primary current taken by the motor versus speed. In both the modes, the current taken by the motor in case of single-layer windings is more as compared to double-layer winding, which results in more power, better power factor, and efficiency.

Table 2 demonstrates that by using changed stator winding, the fundamental factors which specifically identify with the efficiency of the motor can be enhanced even when the motor is working under the evaluated voltage. The general execution of the motor as examined above will rely on the stator and rotor parameters as arranged in Table 2.

The general execution of the motor as examined above will rely on the stator and rotor parameters as arranged in Table 2.

Table 2 demonstrates the design parameters of 5 HP three-phase squirrel cage induction motor. The execution of the motor as discussed above in Figs. 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 are all relying on these designed parameters.

**Fig. 10** Current versus speed**Table 2** Design parameters of 5 HP induction motor

Particulars	Values
Outer diameter	240 mm
Gap length	1.2 mm
Number of poles	4
Number of stator slots	36
Stator bore diameter	122.4 mm
Angle of slot	10.1°
Depth of stator slots	37.09 mm
Width of stator slot bottom	12.81 mm
Width of stator slot opening	3 mm
Height of stator tooth tang	3.2 mm
Angle of stator tooth tang	42.8°
Fillet radius at stator slot bottom	0.72 mm
Fillet radius at stator slot top	0.36 mm
Number of secondary conductors	38
Outer diameter of rotor	120 mm
Shaft diameter	36 mm

(continued)

Table 2 (continued)

Particulars	Values
Bar width	3.55 mm
Bar thickness	15.6 mm
Rotor tooth tang thickness	2.4 mm
Rotor slot opening width	0.88 mm
Rotor teeth	0.22 mm
End ring height	5.03 mm
Upper width of end ring	15.09 mm
Lower width of end ring	15.09 mm

2 Conclusion

From the above discussion, it is concluded that the supply voltage lies under the appraised voltage in the provincial zones, because of reactive power request by the clients. The change of vitality is not much possible from the utility side; anyway, the execution of thousands of Induction motors which are working under evaluated voltages can be enhanced by adjusting the stator and rotor structured at improved value with another mix of stator winding plans. In this investigation article, it is seen that the general execution regarding torque, control created, and efficiency of IM can be enhanced even when these sorts of motors are working under appraised voltage.

References

- Shridhar, L., Singh, B., Jha, C.S., Singh, B.P.: Design of energy efficient motor for irrigation pumps operating under realistic conduction. IEE Roc.-Electrical Power Appl. **141**(6), 269–274 (1994)
- Banerjee, B.: Assessment of voltage unbalance. IEEE Trans. Power Deliv. **16**(4), 782–790 (2001)
- Williams, J.E.: Operation of three-phase induction motors on unbalanced voltages. AIEE Trans. Pt-III-A, Power Appar. Syst. pp. 125–133. (1954)
- Jawad, F., Ebrahimpour, H., Pillay, P.: Influence of unbalanced voltage supply on efficiency of three phase squirrel cage induction motor and economic analysis. Energy Convers. Manag. **47**, 289–302 (2006)
- Hasuike, K.: Efficiency improvement study of low voltage three phase squirrel cage induction motor for general purpose. IEEE Trans. Power Appar. Syst. **pas-102**(12) (1993)
- Pillay, P., Hofmann, P., Manyage, M.: Derating of induction motor operating with combination of unbalanced voltages and over-or-under voltages. IEEE Power Eng. Soc. Winter Meet. **3**, 1365–1371 (2001)
- Gilbert, A.J.: The Influence of Rotating Machine Design Standards on the Design of Traction Supplies, Conference on Main Line Railway Electric Power, pp. 241–245 (1980)
- Wang, Y.J.: Analysis of effects of three-phase voltage unbalance on induction motors with emphasis on the angle of the complex voltage unbalance factor. IEEE Trans. Energy Convers. **16**(3), 270–275 (2001)

9. Emmanuel, B., Aldo, B., Andrea, C.: The incremental design efficiency improvement of commercially manufactured induction motors. *IEEE Trans. Ind. Appl.* **49**(6), 2496–2504 (2013)
10. Yaw, J.W.: An analytical study on steady-state performance of an induction motor connected to unbalanced three-phase voltage. *IEEE Power Eng.* **1**, 159–164 (2000)
11. Sousa Santos, V., Percy, R., Julio, R.: Procedure for determining induction motor efficiency working under distorted grid voltages. *IEEE Trans. Energy Convers.* **30**(1), 331–339 (2015)
12. Gnacinski, P.: Prediction of windings temperature rise in induction motors supplied with distorted voltage. *Energy Convers. Manag.* **49**, 707–717 (2008)
13. Jawad, F., Ebrahimpour, H.: Precise derating of three phase induction motors with unbalanced voltages. *Energy Convers. Manag.* **48**, 2579–2586 (2007)
14. Lee, Ching-Yin: Effects of various unbalanced voltages on the operation performance of an induction motor under the same voltage unbalanced factor conduction. *Electr. Power Syst. Res.* **47**, 153–163 (1998)
15. Saini, R.K., Saini, D.K., Gupta, R., Verma, P.: Optimized design of three phase squirrel cage induction motor based on maximum efficiency operating under the rated voltage – based on software platform. *Indian J. Sci. Technol.* **9**(21), 1–7 (2016). <https://doi.org/10.17485/ijst/2016/v9i21/92002>
16. Rusu, T., Szabo, L., Birte, O., Martis, C.S.: Script controlled modeling of low noise permanent magnet synchronous machines by using JMAG designer. *J. Comput. Sci. Control. Syst.*, 91–94 (2013)
17. Motor models furthering model based Design–Introducing JMAG RT. Ed: JSOL Corp., Error! Hyperlink reference not valid. Date Accessed: 30 Aug 2010
18. Saini, R.K., Saini, D.K., Gupta, R., Verma, P.: Design improvement and assessment of efficiency of three phase induction motor operating under the rated voltage. In: Proceeding of International Conference on Intelligent Communication, Control and Devices. Advances in Intelligent Systems and Computing, vol. 479, https://doi.org/10.1007/978-981-10-1708-7_56
19. Saini, R.K., Saini, D.K., Gupta, R., Verma, P.: Design, analysis, and assessment of efficiency of three-phase squirrel-cage induction motors operating under the rated voltage: a design consideration for rural areas. In: Intelligent Communication, Control and Devices. Advances in Intelligent Systems and Computing, vol. 624. https://doi.org/10.1007/978-981-10-5903-2_2
20. Saini, R.K., Saini, D.K., Gupta, R., Verma, P.: A software approach for the prediction of efficiency of three phase induction motor at optimize iron losses. In: 2016 7th India International Conference on Power Electronics (IICPE). <https://doi.org/10.1109/iicpe.2016.8079353>

A Rate Control Algorithm to Improve TCP over RFID Reader Network



R. Radha, Amit Kumar Tyagi and K. Kathiravan

Abstract In many emerging RFID applications such as location tracking, multiple readers are involved in collecting tag information and sending it to a computer. A large number of computers and readers are augmented into the network for efficient data processing and local decision-making. This type of ad hoc RFID reader network uses TCP as a transport layer protocol. But TCP cannot perform well in any wireless network due to its burst nature of data transmission. This work proposes a rate-based transmission algorithm which is implemented as a layer between TCP and network layer. The algorithm proposed in this work ensures that packets are sent one after another with less delay (i.e., between them). Here, exponential average of an end-to-end delay is used as a metric in determining the delay between the packets. This delay reflects the congestion status in the network and avoids contention between successive data packets. The evaluation of the performance of our algorithm against TCP New Reno using NS 2.35 simulator shows significant performance improvement of throughput, the end-to-end delay, the link layer contentions, and the route failure.

Keywords Location tracking · Rate-based transmission · RFID · TCP

1 Introduction

Radio Frequency Identification (RFID) is a technology for automatic data capture using electromagnetic transmission. It consists of two components: RFID readers and tags [1]. The reader can capture the IDs emitted by RFID tags. They use a

R. Radha (✉)

Department of Computer Science and Engineering, K L University, Hyderabad, Telangana, India
e-mail: radhupriya@yahoo.com

A. K. Tyagi

School of Computer Science and Engineering, Lingayas Vidyapeeth, Faridabad 121002, Haryana, India

K. Kathiravan

Department of Information Technology, Easwari Engineering College, Chennai, Tamilnadu, India

specific frequency and protocol to transmit and receive data. RFID tags are of two types based on the fact whether they are powered by battery. Passive tags do not use battery power. They transmit IDs to the reader by backscattering the part of RF power emitted by reader [1, 2]. Active tags use battery power for transmission of data. RFID application includes factory automation, supply chain, production, smart home environment, health care, parking management, traffic control system, etc. Many applications need to track the physical location of multiple moving objects and make a local decision. The transmission range of tag to reader covers up to a few meters. Many organizations distribute multiple readers over space and connect them to a computer. They augment a number of computers and readers to cover their area of interest. Each computer is involved in huge data processing to eliminate redundancy and communicate with others for optimal local decision-making [3]. Note that this type of network can be formed in an ad hoc manner without the need for any infrastructure.

Basically, TCP (Transmission Control Protocol) is used as a transport layer protocol. It was designed for wired networks. It cannot perform well in wireless network due to its burst data transmission. It misinterprets any type of packet loss that occurs as congestion and reduces the rate of sending drastically. This work proposes a rate-based transmission algorithm, i.e., inserted as a small layer between TCP and network layer (for more details, see Sect. 3). It accepts data packets from TCP but does not send it out immediately. In our proposed approach, packets are delivered out one after the other with sufficient delay between them. An end-to-end delay is used as a metric for calculating this interval. End-to-end delay is measured as the time taken for a packet to reach the destination from source and vice versa. Sufficient interval between the packets helps to avoid contention between successive data packets, and it also reflects the congestion status in the network. This kind of periodic transmission is suitable to the RFID network because reader to tag (passive) communication also happens at a regular interval. Note that we do not make any modifications in the TCP source and destination and add a thin layer at the bottom of TCP and we do not rely on any cross-layer feedback. This algorithm can also be used to improve TCP performance for any application over multi-hop in mobile ad hoc networks.

Hence, remaining of this work is organized as follows: The explanation of the related work is presented in Sect. 2. In Sect. 3, the study of impact of the TCP on the link layer contentions and the system model and protocol architecture are presented in Sect. 4. Further, we discuss the implementation results in Sect. 5. Finally, this work is concluded by providing future work (in brief) in Sect. 6.

2 Related Work

This section discusses various research articles (published in the past decade by several researchers), in which the authors try to control the traffic and reduce the collision over multi-hop Mobile Ad hoc Network (MANET). Fu et al. [4] discussed the reason for having high packet losses, and they found that high competition among

the nodes is the main reason for high packet drops (not buffer overflow, note that primary factor is contributing for maximum packet losses). Also in [4], authors maintain valid congestion Window based on the probability of link layer losses. In [5], Xin proposed an approach with cross-layer design that decrements congestion window based on the network's capability. Round trip delay is measured with queuing time gathered from the nodes present on the way. Further, in [6], Zhang proposed a rate-based transmission that controls congestion window by the measures of medium utilization and rate of contention. Further, in [7], ElRakabawy et al. proposed a method of rate control with out of interference period and variable rate RTT. Later in [8], Sundaresan discussed a method of rate-controlled transmission by using queuing delay. Further, in [9], Ehsan et al. proposed a control (contention) approach where TCP receiver tracks the contention delay and achieves goodput. In [9], the traffic rate is also computed and then updated back to the source to look after the transmission rate. Further, in [10], Jubari et al. determined and controlled the collision between data and acknowledgment packets.

Also in [10], the authors track the packet drop events by tracking the inter-arrival rate among acknowledgments and control the rate of transmission among the data packets. In [11], Jiwei Chen et al. proposed an adaptive ACK scheme (delayed) for ad hoc and hybrid networks. In [12], Tang Lun et al. discussed a Beaconing rate control approach (1-D Markov model). In [12], a beacon rate control scheme mitigates the congestion and maximizes the beacon's delivery efficiency. Further, in [13], Kaixiong Zhou et al. proposed a distributed channel allocation (also a rate control approach) to solve the cross-layer design problem. Further, in [14], Pham Thanh Giang and Kenji Nakagawa also established cooperation among TCP flows with channel access and rate adaptation using cross-layer design. Farzaneh et al. [15] proposed a dynamic TCP-MAC interaction approach to reduce the total number of induced ACKs based on the current channel condition. In last, Carlos De M et al. [16] suggest several mechanisms to enhance TCP performance such as for transmitting forward data and reverse data ACKs to minimize the capture likelihood.

Hence, this section discusses existing work related to our problem. Now, the next section will discuss the effect of TCP on link layer protocol in brief.

3 Effect of TCP on Link Layer Protocol

Transport Control Protocol (TCP) is a transport layer protocol which is designed for wired network, where the losses of packets occur due to congestion. TCP assumes any type of packet loss as congestion and reduces the rate of sending. TCP does not perform well in wireless network due to the inappropriate congestion control algorithm. TCP is burst in nature. It may deliver a group of two or three packets together without worrying about the time interval among them. It is clocked by ACKnowledgment (ACK).

Whenever it receives an acknowledgment, it executes congestion control algorithm and pumps more data into the network, which leads to severe contention at

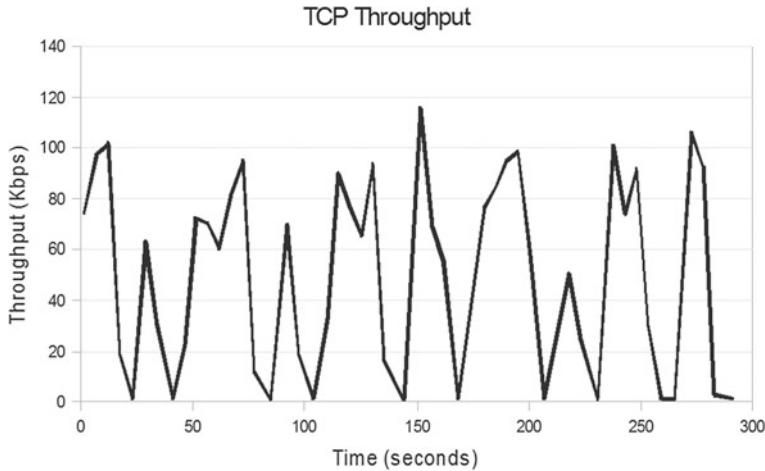


Fig. 1 Variation in TCP throughput

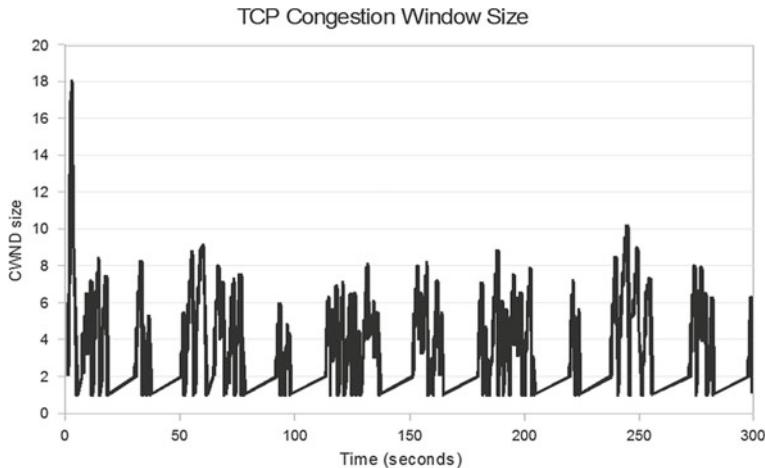


Fig. 2 Variation in TCP congestion window size

MAC layer. The contention results in co-channel interference [8]. This causes collision losses and thereby misinterpreted route failures. It increases the overhead and reduces the throughput finally. Figures 1 and 2 show the variation in TCP throughput and congestion window size for a static horizontal chain of eight nodes, which is shown in Fig. 3 with seven hop TCP connection. Hence, there is a drastic reduction in throughput at many points due to the intra-flow interference. The congestion window is reset to slow start phase at many points.

The contention which occurs at MAC layer can be divided into two types: intra-flow contention and inter-flow contention. The intra-flow contention occurs between

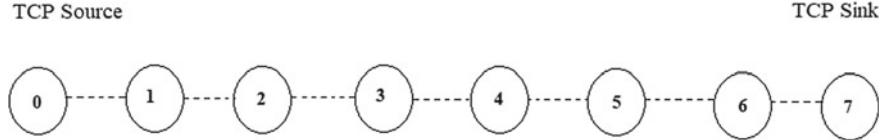


Fig. 3 Horizontal static network

the data packets or between data and acknowledgment packets of the same TCP connection to access the wireless medium whereas the inter-flow contention takes places between the packets of neighboring or parallel TCP connections to access the medium.

The intra-flow contention can be controlled using rate controlled transmission at TCP source. The packets can be sent one after the other with sufficient time intervals. The interval of time between the deliveries of successive packets should be carefully selected such that it does not cause contention and also it should not lead to increased delay which causes unnecessary timeouts for the TCP packets. We explain our rate control algorithm in Sect. 4.

4 System Model

This section discusses the system architecture in Fig. 4. The reader periodically emits the electromagnetic waves and collects tag information which is sent to a computer. Since each computer is connected with multiple readers, it generally receives duplicate tag information. It is the responsibility of the computer to eliminate duplications and communicate with other computers to facilitate local decision-making. The protocol architecture is explained in Fig. 5. Our protocol is implemented as a layer between TCP and network layer. It includes three functions: (a) data processing, (b) rate control, and (c) ACK processing.

Our proposed model can be discussed as follows: It accepts data from TCP layers but delivers them one by one at a specific time interval. This time interval is calculated dynamically by the rate control algorithm. The time interval should be carefully selected, it should be large enough which may otherwise lead to contention, and it should be small which may otherwise cause unnecessary timeouts for the queued TCP packets and reduce the overall throughput. We have selected exponential average of end-to-end delays as the values for calculating the inter-packet delivery period. The end-to-end delay in the forward and reverse path reflects the congestion in the network. Note that the end-to-end delay is smaller for less congested network and high for more congested network. We also prevent contention between the successive packets by setting the inter-packet delivery period approximately to the end delay. We assume that exponential average of end-to-end delay of recent n packets will reflect the end-to-end delay for the new packet to be transmitted. We find end-to-end

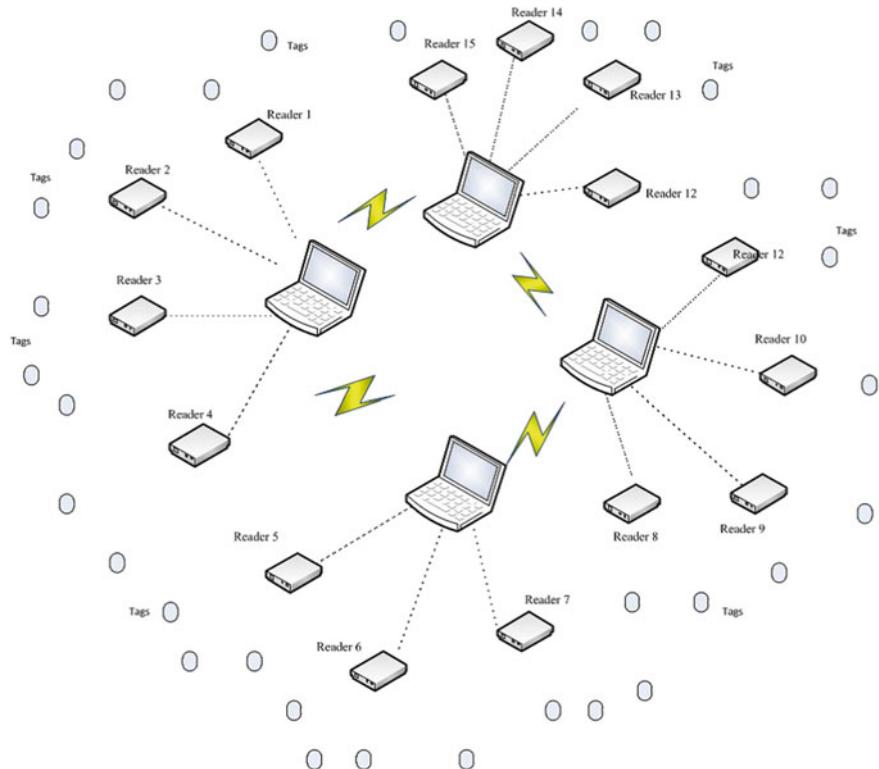


Fig. 4 System architecture

delay for every TCP (forward delay) and corresponding TCP ACK (reverse delay) packet. Here, forward delays are calculated as the interval between the time in which the packet is sent from the thin layer and the time when the packet reaches the destination TCP. And reverse delay is calculated in the reverse manner. The average of forward and reverse delay gives the delay for the recent packet. When the layer receives TCP ACK, it extracts the time stamps stored by the TCP sink and forwards it to the TCP layer immediately. The rate control algorithm updates the inter-packet delivery period using Fig. 5 process.

Here, α is the weightage factor which is given the value of 0.4 in our case, T_{i1} is the time at which TCP Data (i) was sent out by source from the thin layer, and T_{i2} is the time at which TCP Data (i) reaches TCP destination. T_{i4} is the time at which TCP ACK reaches the thin layer in the source. T_{i3} is the time at which TCP ACK was sent out from destination. These time stamps are extracted from the received acknowledgment. E_{if} is the forward delay, E_{ir} is the reverse delay, and $Delay(i)$ is the average of forward and reverse delay of i th transmission. $Exp\ Delay(i+1)$ is the exponential mean of recent i transmissions which will be updated as current

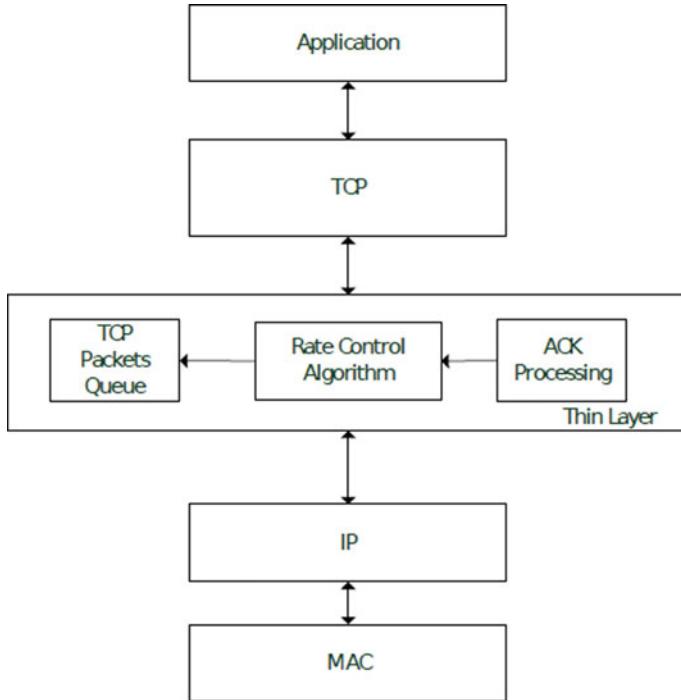


Fig. 5 Proposed protocol architecture

Inter-packet Delivery Period (IPD). This IPD will be updated during the arrival of every acknowledgment.

Whenever the duplicate acknowledgments are received, it does not execute the rate control algorithm, but they are immediately forwarded to the TCP layer. We have used a timer called transmit timer which is set for the duration of period of inter-packet delivery. On the expiry of this, a new packet is sent out.

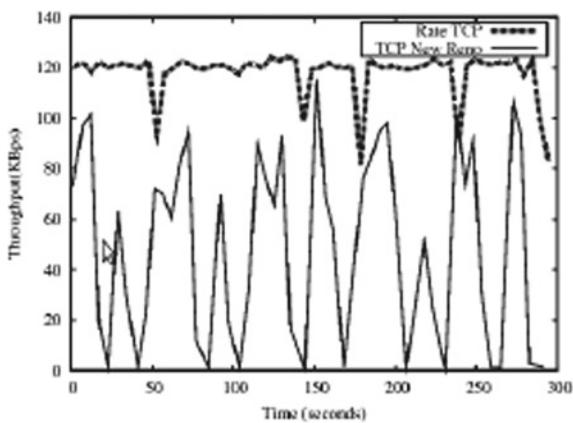
5 Simulation Results

The performance of our solution is evaluated using rate TCP with New Reno using ns2 simulator. We focus a horizontal setup of eight nodes over static multi-hop ad hoc networks. Node 0 as the sender and Node 7 as the receiver. The configuration/parameters of our topology are listed in Table 1.

Table 1 Simulation parameters

Simulation time	300 s
Starting time	At 1.0 s
Number of nodes	8
Source, destination node	0, 7
Routing and MAC protocol	AODV and IEEE 802.11
Transport protocol	Rate TCP/TCP New Reno
Transmission limit	250 m
Interference limit	550 m
Distance among nodes	200 m

Fig. 6 Variation in TCP throughput

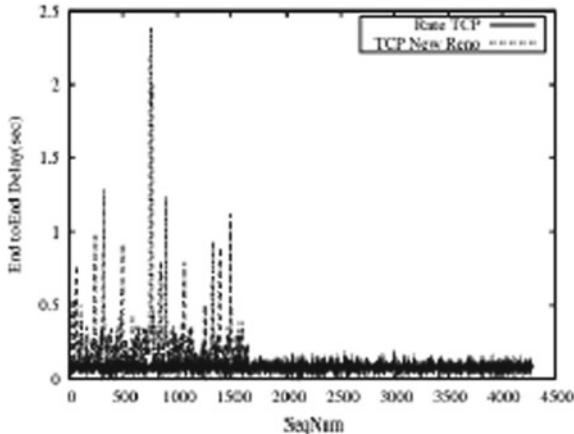


5.1 Throughput

In Fig. 6, Rate TCP is evaluated against TCP New Reno in terms of throughput which is defined as the amount of bytes received by destination per second. The throughput of TCP New Reno is highly variable throughout the simulation. It is due to the timeouts and mis interpreted route failures. Rate TCP performs better than TCP New Reno continuously, and it comes down in few positions because of the collision to the collisions between data and acknowledgments that leads to mis leaded route error followed by throughput reduction.

5.2 End-to-End Delay

Figure 7 shows the path delay for every packet emitted from the source. End-to-end delay is defined as the time taken for the packet to reach the destination after traveling through multiple hops. As packet generation is high for rate TCP, the plot is more

Fig. 7 End-to-end delay

for rate TCP. End-to-end delay is constantly less for Rate TCP than TCP. It shows that the local hop delay is also less for rate TCP.

5.3 Route Failures and Timeouts

Figure 8 shows the performance of rate TCP in a static topology free from mobility. Rate TCP has much less number of mistaken route errors, time exceeds, and retransmissions.

There is no mobility-related losses and note but route failure occurs due to the interference among the packets. Repeated route errors lead to much overhead and scarcity of bandwidth.

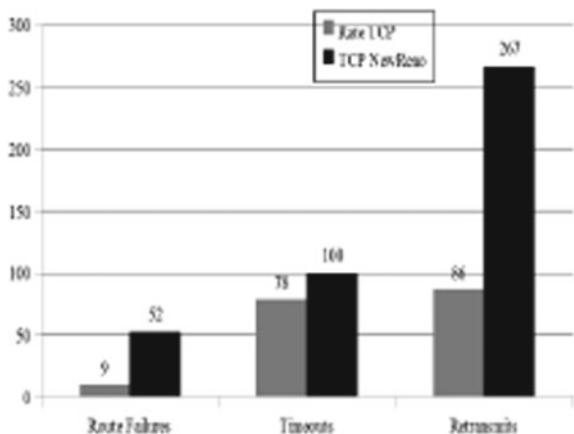
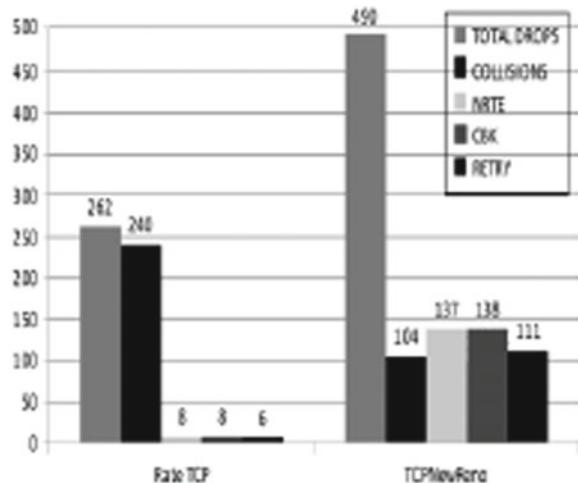
Fig. 8 Route failures, retransmits, and timeouts

Fig. 9 Packet drops

5.4 Packet Drops

Figure 9 describes all types of packet drops. The possible packet drops that may take place are interference losses, NRTE means No Route which is a loss due to unavailability of RTENTRY in routing table to transmit the packet. RETRY is the loss in MAC layer after the maximum attempt of retransmissions. Collision is high while comparing other losses. But the number of drops is very less than NewReno.

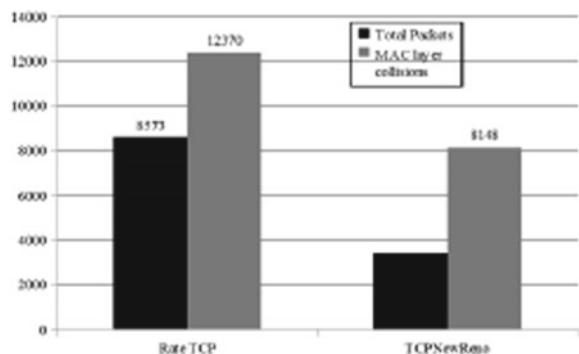
5.5 MAC Layer Collisions

Figure 10 represents the interference drops of MAC layer RTS/CTS packets. The collision also happens while transmitting the RTS. The collision is more in TCP than Rate TCP. Rate TCP had 12370 collisions for transmitting 8573 packets in the transport layer. Hence, in the last next section, we will conclude this work in brief.

6 Conclusion

This work proposed a rate control algorithm that is inserted as a middle ware between TCP and network layer. It accepts data from TCP but delivers them one after the other with sufficient interval among them. This interval is dynamically measured based on congestion that is occurred in the forward and reverse path. The exponential average of the end-to-end delay is considered as inter-packet delivery period. This kind of periodic transmission is suitable to the RFID network because reader to tag (passive)

Fig. 10 MAC layer collisions



communication also happens at a regular interval. The limitation of our algorithm is that it does not consider the spatial reuse property of the wireless medium. This issue/work can be considered as a future work.

References

1. Finkenzeller, K.: RFID Handbook. Wiley, New York (2003)
2. Sarma, S., Engels, D.W.: On the future of RFID tags and protocols. Technical report MIT-AUTOID-TR-018, Auto-ID Center (2003)
3. Vaidya, N., Das, S.R.: RFID-based networks: exploiting diversity and redundancy. SIGMOBILE Mob. Comput. Commun. **1**, 2–14 (2008)
4. Fu, Z., Zerfos, P., Luo, H., Lu, S., Zhang, L., Gerla, M.: The impact of multihop wireless channel on TCP throughput and loss. In: Conference of the IEEE Computer and Communication, vol. 3, pp. 1744–1753 (2003)
5. Zhang, X.M., Zhu, W.B., Li, N.N., Sung, D.K.: TCP Congestion Window Adaptation Through Contention Detection in Ad-Hoc Networks. Vehicular Technology, IEEE Transactions on **9**, 4578–4588 (2010)
6. Zhang, X., Li, N., Zhu, W., Sung, D.K.: CP transmission rate control mechanism based on channel utilization and contention ratio in Ad-hoc networks. IEEE Commun. Lett. **4**, 280–282 (2009)
7. ElRakabawy, S.M., Lindemann, C.: A practical adaptive pacing scheme for TCP in multihop wireless networks. IEEE/ACM Trans. Netw. **4**, 975–988 (2011)
8. Sundaresan, K., Anantharaman, H.-Y.H., Sivakumar, A.R.: ATP: a reliable transport protocol for ad hoc networks. IEEE Trans. Mob. Comput. **6**, 588–603 (2005)
9. Hamadani, E., Rakoccevic, V.: TCP contention control: a cross layer approach to improve TCP performance in multihop AdHoc networks. In: Proceeding software 5th international conference on Wired/Wireless Internet Communications (WWIC’07), pp. 1–16. Springer, Berlin (2007)
10. Jubari, A., Othman, M.: An adaptive delayed acknowledgment strategy to improve TCP performance in multi-hop wireless networks, pp. 0929–6212 (2012)
11. Chen, J., MarioGerla, Y.L., Sanadidi, M.Y.: TCP with delayed ack for wireless networks. Ad Hoc Netw. **7**, 1098–1116 (2008)
12. Lun, T., Yifu, L., Qianbin, C.: Optimized beaconing rate control for vehicular Ad-Hoc networks. J. China Univ. Posts Telecommun. **22**(6), 10–17 (2015)

13. Zhou, K., Gong, C., Nan, W., Zhengyuan, X.: Distributed channel allocation and rate control for hybrid FSO/RF vehicular Ad Hoc networks. *J. Opt. Commun. Netw.* **9**(8), 669–681 (2017)
14. Giang, P.T., Nakagawa, K.: Cooperation between channel access control and TCP rate adaptation in multi-Hop Ad Hoc networks **98.B**, 79–87 (2015)
15. Armaghani, F.R., Jamuar, S.S., Khatun, S., Rasid, M.F.A.: Performance analysis of TCP with delayed acknowledgments in Multi-hop Ad-hoc networks. *Wirel. Pers. Commun.* **4**, 791–811 (2011)
16. De Cordeiro, M., Das, S.R., Agrawal, D.P.: COPAS: dynamic contention-balancing to enhance the performance of TCP over multi-hop wireless networks. In: Proceedings of the 10th International Conference on Computer Communications, and Networks, pp. 382–387 (2002)

A Binary Harmony Search Algorithm Approach for Security Constrained Allocation of PMUs to Obtain Complete Observability



Manam Ravindra, Rayapudi Srinivasa Rao
and Vonteddu Shannmukha NagaRaju

Abstract This paper presents security constrained optimal deployment of Phasor Measurement Units (PMUs) in network to obtain complete observability at minimum cost function. An effective Binary Harmony Search Algorithm (BHSA) is proposed for deployment of PMUs at sensitive buses with an increase in redundancy to provide accurate measurements for state estimation. Sensitive buses are derived from proposed VSI index formulated with load flow solution to provide real-time control for system. Redundancy with and without security constrained PMU location is determined at every bus of network to show effectiveness of security constrained PMU allocation with complete observability. Zero Injection (ZI) bus constraint modeling is considered in optimization to decrease PMU locations. Security constrained PMU location with and without ZI constraint modeling is compared to show its importance with sensitive buses. Complete Network Observability Redundancy Index (CNORI) is proposed to measure the reliability with sensitive constrained deployment of PMUs. IEEE –14, –30, and 57–bus test case systems are considered for MATLAB simulations and outcome results are justified with other proposed methods cited in literature survey.

Keywords Binary harmony search algorithm (BHSA) · Phasor measurement unit (PMU) · Voltage stability index (VSI) · State estimation · Sensitive buses · Zero injection (ZI) buses

Please note that the LNCS Editorial assumes that all authors have used the western naming convention, with given names preceding surnames. This determines the structure of the names in the running heads and the author index.

M. Ravindra · R. S. Rao · V. S. NagaRaju (✉)
Department of Electrical and Electronics Engineering, Jawaharlal Nehru Technological University, Kakinada, India
e-mail: raj.shanmukha@gmail.com

M. Ravindra
e-mail: ravieejntu@gmail.com

R. S. Rao
e-mail: srinivas.jntueee@gmail.com

1 Introduction

In wide area networks, state estimation performed with data available through SCADA systems results in inaccurate states of network. The phasor measurements associated with states of system are proved to be accurate compared to RTU measurements. This leads to the advance of synchrophasor measurement, i.e., phasors associated with states of network are synchronized with time, which is provided by satellite known as Global Position Systems [1]. Deployment of PMUs at all buses in network is unfeasible leading to high economic cost. PMUs should be allocated in a power network in a process that the system would not lose complete observability. Observability should be achieved in such a manner to increase redundancy at all buses in network. The methods for optimization are mainly divided into two methods intelligent search methods and mathematical methods. Among the intelligent-search methods like Binary Particle Swarm Optimization (BPSO) [2], Chemical Reaction Optimization (CRO) [3], and Biogeography-based Optimization (BGO) [4], proposed do not guarantee that global optimal result is obtained. The mathematical methods like Integer Quadratic Programming (IQP) methods [5] are nonlinear where this later formulated to Integer Linear Programming (ILP) in [6]. Gou [7] proposed a generalized integer linear programming method considering the effect ZIB modeling and depth of observability for optimal location of PMUs. In this paper, the security issue is considered as one of the objectives such that sensitive buses are identified to locate PMUs which can avoid blackouts with synchronized communication. VSI is formulated for identification of sensitive buses that are affected by heavy load changes. However, there is a necessity to install PMUs at weak or sensitive buses with complete observability in order to determine accurate measurements at normal and abnormal conditions. BHSA is designed for optimal deployment of PMUs with complete network observability.

2 Problem Formulation

The objective is formulated for optimal deployment of PMUs to minimize cost and redundancy with complete network observability. The cost function formulated as

$$\text{Min} \sum_{p=1}^n C_p x_p \quad (1)$$

$$\text{Subject to } AX \geq B \quad x_p \in (0, 1) \quad (2)$$

where C_p is written as cost coefficient of PMU located at bus p in network, $X = [x_1 \ x_2 \ x_3 \ \dots \ x_n]^T$ is a decision variable matrix, and B is an array of observability constraints, i.e., $[1 \ 1 \ 1 \ 1 \ \dots \ 1]_{n \times 1}^T$. x_p is binary decision variable, and A is bus connectivity matrix which is defined as

$$x_p = \begin{cases} 1 & \text{if PMU is deployed at bus } p \\ 0 & \text{otherwise} \end{cases}$$

$$A_{p,q} = \begin{cases} 1 & \text{if } p = q \text{ or connected to each other} \\ 0 & \text{otherwise} \end{cases}$$

Detection of sensitive buses and deployment of PMUs at sensitive buses giving priority can increase redundancy of every bus and total system with complete observability.

3 Overview of BHS Algorithm

HSA is a meta-heuristic algorithm proposed by Geem et al. [10] and developed by R. S. Rao for distribution networks [11]. This is derived when a group of musicians (population) plays musical instruments or harmony, to get out pleasing music (global best solution). This state is considered as fitness function. The binary optimization problem can be derived as

$$\text{Min } F(x) \quad (3)$$

$$\text{Subjected to } x_i \in (0, 1) \quad i = 1, 2, \dots, N \quad (4)$$

where $F(x)$ is an objective function, x is binary decision variable and N is number of variables.

The solution vector assumed as row vector $[x_1 \ x_2 \ x_3 \ \dots \ x_n]$ of n -bus, generated with random decision variables (0, 1). The row vectors of HMS size number of population which satisfies the subjected constraints are arranged in a matrix initializing binary harmony memory as follows:

$$BHM = \left[\begin{array}{ccccc} x_1^1 & x_2^1 & \dots & x_{n-1}^1 & x_n^1 \\ x_1^2 & x_2^2 & \dots & x_{n-1}^2 & x_n^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1^{HMS-1} & x_2^{HMS-1} & \dots & x_{n-1}^{HMS-1} & x_n^{HMS-1} \\ x_1^{HMS} & x_2^{HMS} & \dots & x_{n-1}^{HMS} & x_n^{HMS} \end{array} \right] \quad (5)$$

The objective is fitness function that satisfies constraints. The weight vector is used to define cost value depending on factor of installation and manufacture criteria:

$$\text{fitness}(x) = \sum_{p=1}^n C_p x_p \quad (6)$$

where weight matrix (C_p) in form of a diagonal matrix, normally considered as a diagonal unit vector or weight, i.e., increased or decreased.

A novel harmony vector $[\hat{x}_1 \hat{x}_2 \dots \hat{x}_n]$ comprising decision variables (0, 1) is developed built on memory, pitch adjustment, and randomly chosen decision variables:

```

if rand(0, 1) < HMCR
     $\hat{x}_p \leftarrow \hat{x}_p \in [x_p^1, x_p^2, \dots x_p^{HMS}]$ 
else
     $\hat{x}_p \leftarrow \hat{x}_p \in randi([0, 1])$ 
end

```

where $rand(0, 1)$ is MATLAB term defined to select the random value between 0 and 1 value, and $randi([0, 1])$ term selects randomly 0 or 1 value for the solution vector. The parameter variables are inspected with adjustment of pitch rate as follows:

```

if rand(0, 1) < PAR
     $\hat{x}_p = \hat{x}_p + randi([0, 1])$ 
else
     $\hat{x}_p = \hat{x}_p$ 
end

```

If new harmony $\vec{x} = [\hat{x}_1 \hat{x}_2 \dots \hat{x}_n]$ proved to provide improved fitness value than weak harmony, the new harmony is placed in HM and weak harmony is removed from HM with merge process and evaluation of fitness function to each array vector. The algorithm is run with number of iterations considered and terminated when it is met with optimal output otherwise steps 3 and 4 are repeated.

4 Application of Security Constrained PMU Deployment with BHS Algorithm

4.1 Voltage Stability Index (VSI) for Sensitive Buses

Proper identification of sensitive buses and allocation of PMUs at sensitive buses give accurate states of network such that proper security is provided [12]. The VSI is formulated to find voltage sensitive buses in the network as follows:

$$V_{mean} = \frac{1}{M} \sum_{i=1}^M V(i) \quad (7)$$

Table 1 Voltage stability index for 14-bus system

Bus no.	B1	B2	B3	B4	B5	B6	B7
VSI	0	0	0	0.0061	0.0058	0	0.0059
Bus no.	B8	B9	B10	B11	B12	B13	B14
VSI	0	0.0090	0.0089	0.0052	0.0034	0.0048	0.0111

$$VSI = \frac{V_{mean} - V_0}{V_0} \quad (8)$$

The highest value of indices obtained is most sensitive bus to which the PMU is to be allocated. Consider 14-bus network to which VSI formulation is applied. From Table 1, bus-9 and bus-14 are highest sensitive buses.

4.2 Security Constrained Deployment of PMUs Using BHS Algorithm

The security constrained approach is formulated for the deployment of PMU to optimize redundancy at all buses with complete observability. The cost function is formulated subjected to security constraints and observability constraints for the deployment of PMUs as

$$\text{Min} \sum_{p=1}^n C_p x_p \quad (9)$$

$$\text{Subject to } AX \geq B \text{ and } A_{eq}X_{eq} = B_{eq} \quad (10)$$

where A_{eq} is an array of sensitive buses whose entries are all equal one of order $n \times n$, X_{eq} is an array of decision variable matrix of order $n \times 1$ in which selected buses values equal to one, and B and B_{eq} are an array of observability constraints which can be written as $[1 \ 1 \ 1 \ 1 \ \dots \ 1]_{n \times 1}^T$.

The security bus constrained binary harmony memory is initialized after it satisfies complete observability condition. Table 2 shows the parameters considered in BHS Algorithm. The procedure for optimal allocation of PMU is described through flowchart of BHSA as shown in Fig. 1. Table 3 shows sensitive buses and ZI buses of the different test case systems.

Fig. 1 Flowchart of binary harmony search algorithm

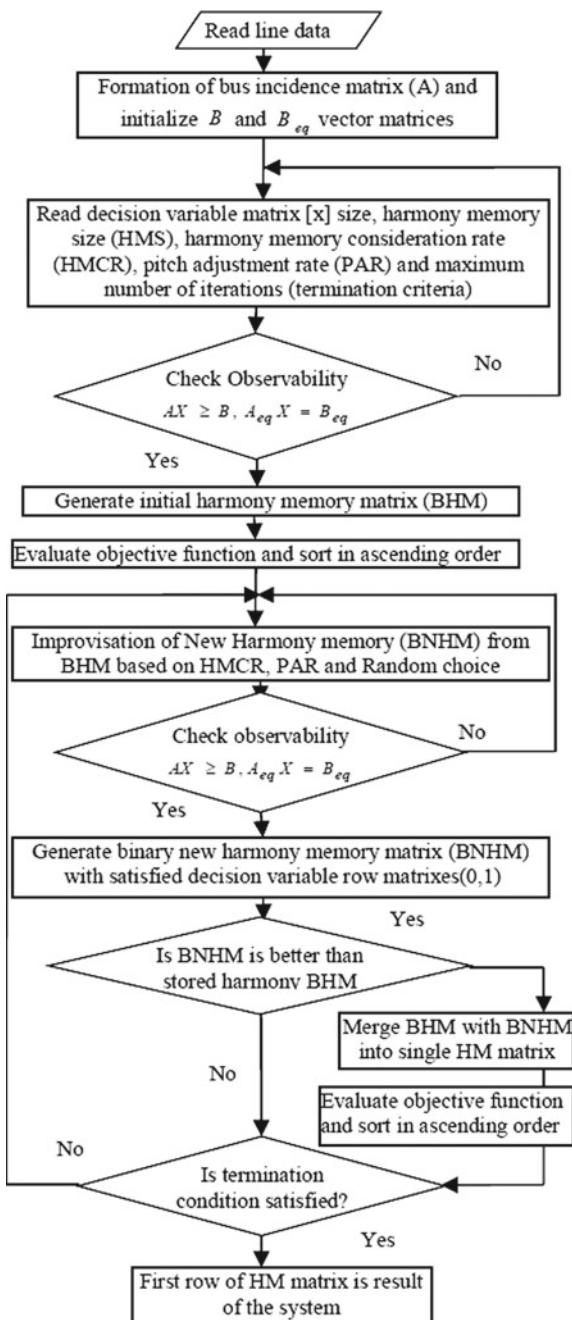


Table 2 BHSA parameters

HMS	30
HMCR	0.95
PAR	0.0001
Maximum iterations	300

Table 3 Sensitive buses

IEEE test systems	Sensitive buses	ZIB buses
14 bus	9, 14	7
30 bus	30, 26, 24	6, 9, 22, 25, 27, 28
57 bus	31, 33, 29, 32	4, 7, 11, 21, 22, 24, 26, 34, 36, 37, 39, 40, 45, 46, 48

4.3 Security Constrained PMU Allocation with ZIB Modeling

For allocation of PMU with minimum cost, the cost function is subjected to constraints to place PMUs at sensitive buses, observability constraints and ZI bus constraints are as follows:

$$\text{Min} \sum_{p=1}^n C_p x_p \quad (11)$$

$$\text{Subject to } AX \geq B, A_{eq}X = B_{eq} \text{ and } \hat{A}X_{zq} = 0 \quad (12)$$

where \hat{A} is an array of ZI buses that are all equal to one. X_{zq} is an array of decision variable matrix $n \times 1$ in which selected buses value is zero.

4.4 Complete Network Observability Redundancy Index (CNORI)

Observability index is formulated to measure reliability condition of network. Observability of complete network is subjected to on redundancy measurement at all buses in network. Bus Observability Index (BOI) and NORI are derived in [6] and are formulated as

$$\mathfrak{N}_p \leq \chi_p + 1 \quad (13)$$

BOI is restricted to maximum incident branches (χ_p) plus one [6]. For bus p , $BOI(\mathfrak{N})$ gives the count of PMUs allocated to measure the performance of bus. For complete network, the CNORI is calculated as

$$CNORI = \sum_{p=1}^n \mathfrak{N}_p \quad (14)$$

5 Results and Analysis

The optimal PMU allocation problem with BHSA is programmed in MATLAB, and it runs on Intel(R) core(TM), the i3 processor at 2.20 GHz with 4 GB of RAM. Figure 2 presents 14-bus network with PMU allocation. To allocate PMUs and check complete observability, different test cases such as 14-, 30-, and 57-bus test systems are considered to analyze the applicability of proposed BHSA. The sensitive buses obtained from VSI proposed are 14 and 9 buses. When we compare 14 and 9 buses, it can be observed that bus-9 is incident with more branches including bus-14. So in this work, bus-9 is considered as the most critical bus as it is much affected with heavy load and high stress due to more branches incident to it.

Sensitive buses are arranged in the table based on the priority of locations. The sensitive buses obtained from VSI proposed are 14 and 9 buses. When we compare 14 and 9 buses, it can be observed that bus-9 is incident with more branches including

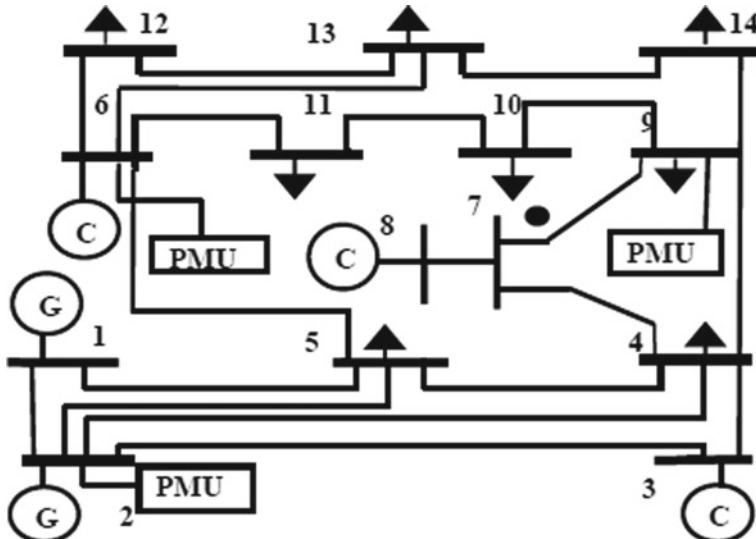
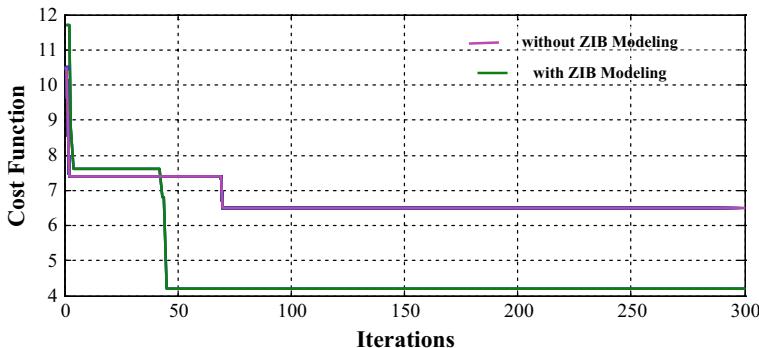


Fig. 2 14-bus system with PMU allocation

Table 4 Security constrained PMU allocations

IEEE test systems	Without ZIB modeling		With ZIB modeling	
	No PMUs	PMU allocations	No PMUs	PMU allocations
14 bus	5	2, 6, 8, 9, 14	3	2, 6, 9
30 bus	10	3, 6, 7, 9, 10, 12, 19, 24, 26, 30	8	3, 7, 10, 11, 24, 26, 29, 30
57 bus	19	2, 4, 5, 10, 15, 19, 20, 29, 30, 31, 32, 33, 40, 41, 45, 49, 50, 51, 56	14	1, 6, 13, 19, 25, 27, 29, 31, 32, 33, 41, 47, 51, 54

**Fig. 3** Comparison of convergences characteristics of 14-bus network

bus-14. Sensitive buses are utilized to allocate PMUs at their locations, and ZI buses are used for further minimize PMUs in network. Table 4 shows a number of PMUs and locations of security bus constrained PMU allocation with and without application of ZIB modeling.

From comparison of Table 4, it is observed that minimum PMUs are allocated at sensitive buses giving major priority for allocation with complete observability. Figure 3 shows a comparison of convergences characteristics of security constrained BHS algorithm with and without ZIB modeling. To decrease PMUs number for allocation, ZIB modeling with sensitive constraints is implemented for security criteria. From this figure, it can be observed that the cost function obtained for 14 bus system is very less, i.e., 4.2 p.u with minimum PMUs at 2, 6, and 9 locations. Figure 4 shows convergence characteristics of BHS algorithm considering sensitive buses with and without ZIB modeling for 30 bus system. CNORI with ZIB modeling attains optimal redundancy, i.e., more than bus number with complete observability. Figure 6 shows the minimum cost function value required for different bus network for PMUs allocation (Fig. 5 and Table 5).

Fig. 4 30-bus network convergence characteristics with sensitive buses

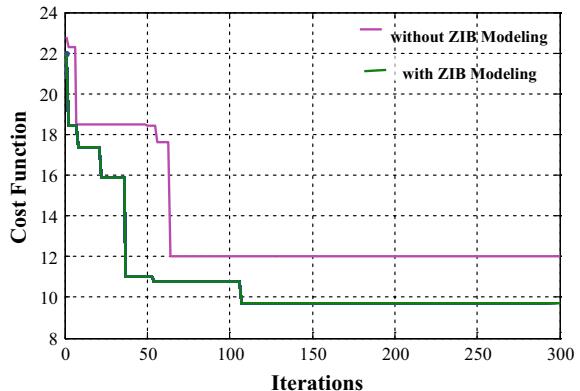


Fig. 5 57-bus network convergence characteristics with sensitive buses

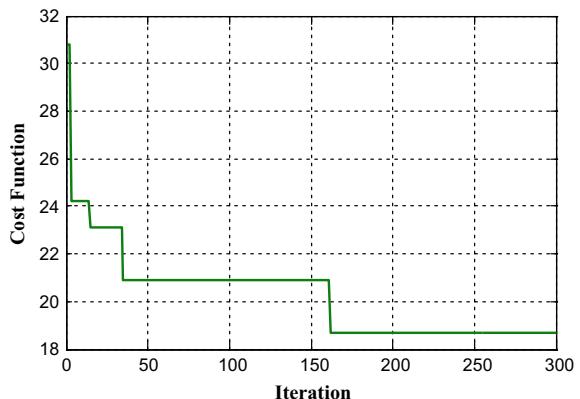


Fig. 6 Cost function comparison of BHSA algorithm with other methods

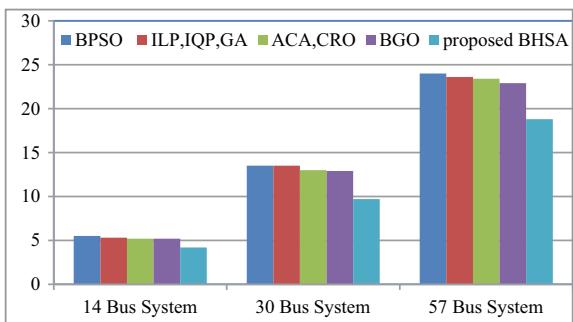


Table 5 Complete network observability redundancy index (CNORI)

IEEE test system	CNORI with ZIB	CNORI without ZIB
14 bus	16	22
30 bus	35	43
57 bus	68	79

6 Conclusion

A new BHS algorithm is presented for optimal allocation of PMUs at sensitive buses with major priority. VSI is formulated considering load increase up to 50%, through which sensitive buses are generated. ZIB modeling is proposed for minimization of PMUs that reduces cost and installation of PMUs in network. The convergence characteristics of security constrained PMU allocation with ZIB modeling and without ZIB modeling are presented to show effectiveness of ZIB modeling. The ZIB modeling shows impact on reduction of PMUs thereby reducing the cost of PMUs required. CNORI was utilized to check performance and increase of redundancy with security constrained allocation of PMUs with complete observability. The BHS algorithm is compared with standard methods to show its effectiveness.

References

1. Phadke, A.G., Thorp, J.S., Karimi, K.J.: State Estimation with phasor measurements. *IEEE Trans. Power Syst.* **1**(1), 233–238 (1986)
2. Ahmadi, A., Alinejad-Beromi, Y., Moradi, M.: Optimal PMU placement for power system observability using binary particle swarm optimization and considering measurement redundancy. *Expert Syst. Appl.* **38**(6), 7263–7269 (2011)
3. Xu, J., Wen, M.H., Li, V.O., Leung, K.C.: Optimal PMU placement for wide-area monitoring using chemical reaction optimization. In: IEEE PES Conference, Innovative Smart Grid Technologies (ISGT), pp. 1–6 (2013)
4. Jamuna, K., Swarup, K.S.: Multi-objective biogeography based optimization for optimal PMU placement. *Appl. Soft Comput.* **12**(5), 1503–1510 (2012)
5. Chakrabarti, S., Kyriakides, E., Eliades, D.G.: Placement of synchronized measurements for power system observability. *IEEE Trans. Power Deliv.* **24**(1), 12–19 (2009)
6. Dua, D., Dambhare, S., Gajbhiye, R.K., Soman, S.A.: Optimal multistage scheduling of PMU placement: An ILP approach. *IEEE Trans. Power Deliv.* **23**(4), 1812–1820 (2008)
7. Gou, B.: Generalized integer linear programming formulation for optimal PMU placement. *IEEE Trans. Power Syst.* **23**(3), 1099–1104 (2008)
8. Müller, H.H., Castro, C.: A Genetic algorithm-based phasor measurement unit placement method considering observability and security criteria. *IET Gener., Transm. Distrib.* **10**(1), 270–280 (2016)
9. Bian, X., Qiu, J.: Adaptive clonal algorithm and its application for optimal PMU placement. In: International Conference IEEE. Communications, Circuits and Systems. vol. 3, pp. 2102–2106 (2006)
10. Geem, Z.W., Kim, J.H., Loganathan, G.V.: A new heuristic optimization algorithm: harmony search. *Simulation* **76**(2), 60–68 (2001)
11. Rao, R.S., Narasimham, S.V.L., Raju, M.R., Rao, A.S.: Optimal network reconfiguration of large-scale distribution system using harmony search algorithm. *IEEE Trans. Power Syst.* **26**(3), 1080–1088 (2011)
12. Ravindra, M., Rao, R.S.: Sensitive constrained optimal PMU allocation with complete observability for state estimation solution. *Eng. Technol. Appl. Sci. Res.* **7**(6), 2240–2250 (2017)

Ledger-Based Sorting Algorithm



Nakshatra Kothi, Pradeep Laxkar, Anuj Jain and Prasun Chakrabarti

Abstract Research on sorting algorithms is a very core research field in computer science. Today, any advancement in sorting algorithms would lead to less time and space required to perform the sorting task. This paper discusses a newly proposed algorithm—‘ledger-based sorting algorithm’. It is basically an advancement to ‘counting sort’, but not exactly counting sort. The purpose of this paper is to reduce the time and space required for sorting when the numbers given to you belong to a particular range (0–n) and all of them are distinct. This paper also compares the newly proposed algorithm with insertion sort, bubble sort, selection sort, quick sort, merge sort and mainly counting sort through graphical time analysis of these algorithms. The paper especially compares counting sort algorithm with the proposed algorithm in terms of both time and space complexity.

Keywords Insertion sort · Bubble sort · Selection sort · Quick sort · Merge sort · Counting sort · Graphical time analysis

1 Introduction

Sorting is the process/technique of arranging elements (records) in ascending/descending order. It is an essential process which is required frequently in computer programming [1]. Sorting is considered to be one of the most rudimentary problems in the ‘analysis and design of algorithms’ by many computer scientists. There are several reasons for it. Sometimes, a task primarily requires sorted information. Algorithms very frequently use sorting as a prime subroutine. Selecting from among a wide diversification of sorting algorithms, and they have their respective techniques. Sorting is a technique/method of historical interest. Many engineering tasks/problems come to the fore by the implementation of sorting algorithms [4].

N. Kothi (✉) · P. Laxkar · A. Jain · P. Chakrabarti
Computer Science Dept, Institute of Technology and Management Universe, Vadodara, Gujarat, India
e-mail: nakshatra.kothicse10@gmail.com

Some of the sorting algorithms taught to a beginner are insertion sort, bubble sort, selection sort and counting sort. These algorithms aid the learner to understand the process and technique of sorting [2–5]. Sorting algorithms are classified on the basis of various criteria. On the basis of memory, sorting is ‘internal’ if it is done in main memory and ‘external’ if it is done in secondary memory. Sorting is also classified as ‘comparison-based sorting’ (bubble sort, insertion sort, merge sort, selection sort, quick sort) and ‘non-comparison-based sorting’ (radix sort, counting sort) on the basis of whether comparison is done or not. Sorting is ‘stable’ if the order of repetitive elements (or records) remains same before and after the sorting [4]. Bubble sort is a customary, but very amateurish, sorting algorithm. This algorithm works by time after time interchanging elements that are proximate to each other in an array that are out of commission (not in working order) [1–5]. The worst case and best case time complexity of bubble sort is $O(n^2)$. In selection sort, selection of an element to be minimum/maximum is done. Then an element which is less/greater than the selected element is replaced with the selected element. The process is repeated for the entire array until it gets sorted [1–4]. The worst case and best case time complexity of selection sort is $O(n^2)$. In insertion sort, an element is selected and a check occurs whether the previous elements are smaller/greater than the selected element. If it is not so, then the selected element will be inserted at the appropriate position [1–5]. The worst case time complexity of insertion sort is $O(n^2)$ and the best case time complexity is $O(n)$. Counting sort is used when numbers in an array are repetitive and in range 0–k, where k < n and n is the number of elements. All the numbers should be positive [3, 4]. The time complexity of counting sort (for best case as well as worst case) is $O(n)$. Merge sort is a divide and conquer-based sorting technique. In this, the problem is divided into two parts. It uses the concept of insertion sort [1, 3, 4]. The time complexity of merge sort (for best case as well as worst case) is $O(n\log n)$. Quick sort is a divide and conquer-based sorting technique. It uses the concept of bubble sort. The worst case time complexity of quick sort is $O(n^2)$ and the best case time complexity is $O(n\log n)$.

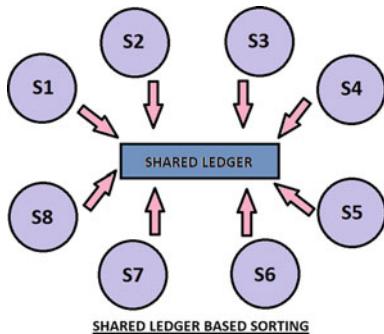
2 Ledger-Based Sorting Algorithm

The proposed algorithm in this paper is basically an advancement to counting sort, but not exactly counting sort. The condition for this algorithm to be applied is that all the numbers should be positive and distinct and belong to a particular range (0–n).

The motivation to develop this algorithm came from a very common activity conducted in any classroom which is the activity of taking attendance of students.

Sometimes when the teacher is in a hurry, he asks the students to call out their roll numbers loudly in appropriate order. And by this, the least roll number student call out his roll number and then succeeding students call out their roll numbers hearing to the most recent roll number spoken. And hence, by this, the teacher is able to note down the roll numbers of students present in ascending order. This process was obviously faster than if the teacher would have taken the attendance by calling out

Fig. 1 Diagram for ledger-based sorting process



the roll numbers and then students responding. Now, what could be a more efficient way of sorting such numbers.

Observe that in this situation, the roll numbers allocated to the students are all distinct and fall in a particular range.

A common ledger could be shared to all the students on the portal of the college which consists of the entire range of roll numbers written in ascending order and then the students can access this ledger and put a present mark on their corresponding number. Finally, when all the students have completed the process you would get all the present students roll number in ascending order. Figure 1 shows a diagram explaining the system.

ALGORITHM:

LedgerSort($\text{Arr}[], n1, n2$) execute	No. of instructions	Cost to
{		
1. Create a ledger array of size $n1$;	1	X1
2. Initialize ledger array by 0;	1	X2
3. for($\text{int } j=1; j \leq n2; j++$)	$(n2+1)$	X3
{		
4. $\text{ledger}[\text{Arr}[j]] = 1;$	$n2$	X4
}		
5. for($\text{int } i=1; i \leq n1; i++$)	$(n1+1)$	X5
{		
6. if($\text{ledger}[i] == 1$)	$n1$	X6
{		
7. $\text{Arr}[i] = i;$	$n1 * d_i$	X7
}		
}		
}		

Here, $\text{Arr}[]$ is the array containing the data of present roll numbers. The size of array $\text{Arr}[]$ is n_2 . Hence, n_2 is the total number of students present. Size of ledger array is n_1 . Hence n_1 is the total number of students. Each element of array $\text{Arr}[]$ can be assumed as a student accessing the ledger array and making an entry of 1 at his corresponding roll number. Finally, the indices of the ledger array with data value 1, which are nothing but the present roll numbers, are stored in the original array. They will definitely get stored in ascending order as the ledger indices are, of course, already sorted as it is the property of any array.

So, the time complexity $T(n)$ of the proposed algorithm is as follows.

$$T(n) = X_1 + X_2 + (n_2+1)*X_3 + n_2*X_4 + (n_1+1)*X_5 + n_1*X_6 + n_1*d_i*X_7$$

Now, for worst case $d_i = 1$, so

$$T(n) = (X_5+X_6+X_7)*n_1 + (X_3+X_4)*n_2 + X_1 + X_2 + X_3 + X_5$$

Let, $X_5+X_6+X_7 = A$, $X_3+X_4 = B$, $X_1 + X_2 + X_3 + X_5 = C$

$$\text{So, } T(n) = A*n_1 + B*n_2 + C = O(n_1) \quad [\text{Since } n_1 > n_2] = O(n) \quad [\text{Let } n_1 = n]$$

Now, for best case $d_i = 0$, so

$$T(n) = (X_5+X_6)*n_1 + (X_3+X_4)*n_2 + X_1 + X_2 + X_3 + X_5$$

Let, $X_5+X_6 = A$, $X_3+X_4 = B$, $X_1 + X_2 + X_3 + X_5 = C$

$$\text{So, } T(n) = A*n_1 + B*n_2 + C = O(n_1) \quad [\text{Since } n_1 > n_2] = O(n) \quad [\text{Let } n_1 = n]$$

We can analyse that the algorithm put forward in the paper is obviously quicker than insertion sort, bubble sort and selection sort whose worst case time complexities are of the order $O(n^2)$ since the worst case time complexity of the proposed algorithm is of the order $O(n)$.

Also, the ledger algorithm is obviously speedier than merge sort and quick sort whose best case time complexities are of the order $O(n\log n)$ since the worst case time complexity of the proposed algorithm is of the order $O(n)$.

Hence, we see that the time complexity of the ‘ledger-based sorting algorithm’ is same as that of counting sort algorithm. But counting sort algorithm has 4 ‘for’ loops and ‘ledger-based sorting algorithm’ has only 2 ‘for’ loops and also it does not require any extra array to store the sorted numbers. It stores the sorted numbers in the same array. So the execution time of the proposed algorithm will be less as compared to the counting sort algorithm.

Let us perform the space complexity analysis of counting sort algorithm and ‘ledger-based sorting algorithm’.

Now, space complexity is the space required by an algorithm to perform a computational task. Space complexity $S_{(p)}$ is given by

$$S_{(p)} = C_{(p)} + V_{(p)},$$

where $C_{(p)}$ is the space required by constant data; $V_{(p)}$ is the space required by variable data.

Now, for ledger-based sorting algorithm, the space complexity is given as follows. Here, constant data are n_1 , n_2 , i , j , and variable data are A and ledger. Hence, space complexity $S_{(p)}$ is given by

$$S_{(p)} = C_{(p)} + V_{(p)} = 4 + n_2 + n_1.$$

Now, for counting sort algorithm and ledger-based sorting algorithm, the space complexity is given as follows.

```
CountSort(Arr[],n,k)
{
1.   for(l=0 ; l <= k ; l++)
      {
2.       C[l] = 0;
      }
3.   for(l=1 ; l <= n ; l++)
      {
4.       C[Arr[l]] = C[Arr[l]] + 1;
      }
5.   for(l=1 ; l <= k ; l++)
      {
6.       C[l] = C[l-1] + C[l];
      }
7.   for(l=n ; l >= 1 ; l--)
      {
8.       B[C[Arr[l]]] = Arr[l];
9.       C[Arr[l]] = C[Arr[l]] - 1;
      }
}
```

Here, constant data are n , k , l , and variable data are Arr , B and C . For constant we take 1 as space complexity and for variable we take the size accordingly.

Hence, space complexity $S_{(p)}$ is given by

$$S_{(p)} = C_{(p)} + V_{(p)} = 3 + (k + 1) + n + k + n = 4 + 2n + 2k.$$

We see that the space complexity of the ‘ledger-based sorting algorithm’ is less than that of the counting sort algorithm. This means that the proposed algorithm is more efficient than the counting sort algorithm with respect to space complexity criteria.

Table 1 The execution time in micro-seconds of bubble sort, selection sort, insertion sort, merge sort, quick sort, counting sort and ledger sort (the proposed algorithm) for the sorting of arrays with increasing number of entries

No. of inputs	Bubble sort	Selection sort	Insertion sort	Merge sort	Quick sort	Count sort	Ledger sort
100	1453	1378	1190	345	321	23	12
600	2293	1986	1231	456	479	34	21
1100	3479	2978	1986	657	623	37	25
1600	6791	5981	4561	1298	1127	43	39
2100	12347	11978	15620	1297	983	29	17
2600	15621	15622	12341	1329	1321	234	124
3100	15622	15623	11090	1346	989	789	341
3600	31206	15622	13467	5479	13590	1023	763
4100	46864	15656	14591	3211	3398	2347	1298
4600	46898	15604	15604	2349	2763	1982	674
5100	46864	31242	15622	1891	1764	973	129
5600	62523	46836	15649	4562	4671	3040	1986
6100	78106	46864	15622	7892	15626	831	679
6600	78107	62486	15621	1993	1298	542	87
7100	109350	46861	31242	3469	15629	979	452
7600	124970	62486	31279	2153	1980	721	327
8100	140592	78107	46862	2289	2147	479	86
8600	171835	78106	46865	2299	2199	631	341
9100	187457	93762	46865	2301	2259	1792	725
9600	203079	109344	46868	2309	2456	1704	839
10100	234282	109350	62486	3481	3289	2398	1237
10600	249942	140594	62486	2389	1059	919	321
11100	296800	140592	78108	2398	1798	439	169
11600	328055	156205	78109	2419	2561	314	87
12100	343668	171830	93731	2427	2677	627	351
12600	374946	187458	78106	2456	4367	587	271
13100	406153	203114	93727	2489	9876	1021	842
13600	437396	218702	93727	2491	2954	1029	900
14100	453055	234283	109348	2516	3247	1297	913
14600	515503	249941	124969	2524	1376	996	812
15100	562367	265562	140593	2317	1024	769	313
15600	577953	281185	140625	2441	997	893	391
16100	624852	296769	156214	3471	3331	1004	787
16600	687371	312427	171836	3267	3728	1001	402
17100	734201	343670	171869	3391	3197	973	123

(continued)

Table 1 (continued)

No. of inputs	Bubble sort	Selection sort	Insertion sort	Merge sort	Quick sort	Count sort	Ledger sort
17600	781102	359289	171836	3691	3744	2791	1761
18100	812345	374875	171836	4198	3344	1011	129
18600	859175	406118	187493	5731	6441	4375	85
19100	890412	421780	187418	4398	5448	997	481
19600	937282	453022	187417	4497	5698	996	367
20100	984175	468643	203041	15622	16793	1004	837
20600	1046664	484224	218698	5691	6731	2012	1296
21100	1077837	515502	218736	2371	2591	657	97
21600	1124697	531125	249975	3287	2893	513	98
22100	1187226	562361	249975	7891	15624	996	451
22600	1218497	593576	265561	2590	2894	1117	998
23100	1296569	609234	281181	3721	3277	1291	614
23600	1343433	640475	296841	2301	3489	999	312
24100	1405919	671716	296769	1984	1999	609	234
24600	1468368	703006	312414	2769	1982	1023	995
25100	1530856	718586	374940	4521	5741	2137	834
25600	1593410	749828	343662	3191	15620	1015	996
26100	1671446	781101	359291	3376	15621	2129	986
26600	1749587	812308	374913	15622	11290	998	353
27100	1812106	827936	406147	6790	18058	649	167
27600	1890214	874759	421776	15722	997	994	223
28100	1983873	906037	437395	16814	17891	781	543
28600	2046399	937238	453054	7569	9453	997	261
29100	2140121	952900	484260	8834	8899	1278	999
29600	2233848	999763	515503	9879	8752	1640	793

3 Result and Discussion

Table 1 shows the output of the C program which shows the execution time in microseconds of insertion sort, bubble sort, selection sort, counting sort, merge sort, quick sort and ledger sort (the proposed algorithm).

Figure 2 represents the graphical time analysis of the proposed algorithm with insertion sort, bubble sort and selection sort. Figure 3 represents the graphical time analysis of the proposed algorithm with merge sort, quick sort and counting sort algorithm. Figure 4 represents the graphical time analysis of the proposed algorithm with counting sort algorithm.

Fig. 2 Graphical time analysis of the proposed algorithm with bubble sort, selection sort and insertion sort algorithm

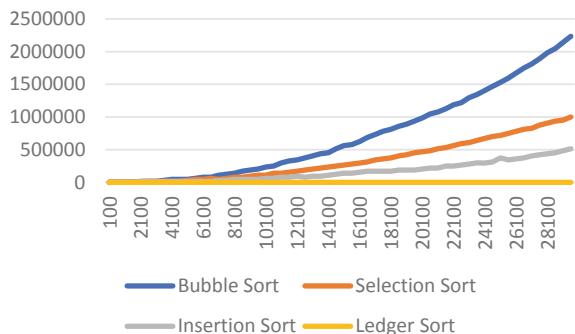


Fig. 3 Graphical time analysis of the proposed algorithm with merge sort, quick sort, and counting sort algorithm

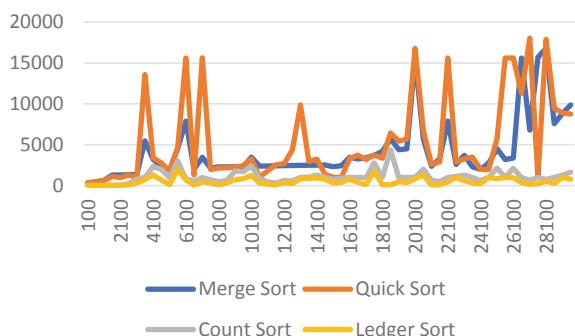
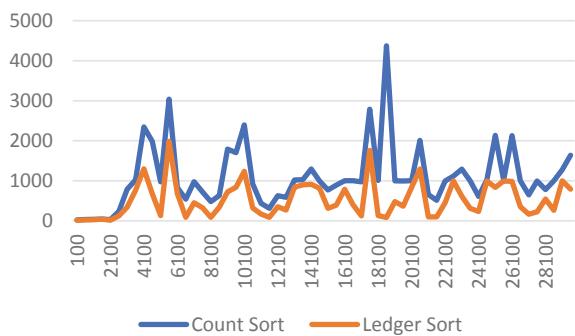


Fig. 4 Graphical time analysis of the proposed algorithm with counting sort algorithm



For each of the above graphs, the x-axis represents the number of inputs to the sorting algorithm, and the y-axis represents the time required in micro-seconds to perform sorting. Table 2 shows the comparison of time complexity of various algorithms with the proposed algorithm. Bubble sort, selection sort and insertion sort are much slower than counting sort and ledger sort as their worst case time complexities (of order $O(n^2)$) are much more than the time complexities of counting sort and ledger sort (of the order $O(n)$). Merge sort and quick sort are faster than insertion, bubble and selection sort but slower than counting sort and the proposed algorithm. Although the time complexity of counting sort and ledger sort is the same, we saw by

Table 2 The comparison of time complexity of various algorithms with the proposed algorithm

Algorithm	Worst case	Best case
Bubble sort	$O(n^2)$	$O(n^2)$
Selection sort	$O(n^2)$	$O(n^2)$
Insertion sort	$O(n^2)$	$O(n)$
Merge sort	$O(n \log n)$	$O(n \log n)$
Quick sort	$O(n^2)$	$O(n \log n)$
Counting sort	$O(n)$	$O(n)$
Ledger sort	$O(n)$	$O(n)$

Table 3 The comparison of space complexity of counting sort and ledger sort algorithm

Algorithm	Space complexity
Counting sort	$4 + 2n + 2k$
Ledger sort	$4 + n^2 + n1$

performing the experiment that the time required for executing is less of ledger sort as compared to counting sort as counting sort algorithm has the presence of more loops than that in ledger sort. Table 3 shows the comparison of space complexity of counting sort and ledger sort since their time complexities were found to be the same. It is clear that ledger sort is much more efficient than counting sort in terms of space complexity.

4 Conclusion

From the result of the experiment, we can conclude that the proposed algorithm, i.e. ‘ledger-based sorting algorithm’, is efficient than insertion sort, bubble sort, selection sort, merge sort, quick sort and especially counting sort in terms of both space complexity and time complexity as it is evident from the graphical time analysis shown in Fig. 2, 3 and 4. So, we can conclude that for a particular condition, i.e. the given number belonging to a particular range (0–n) and all of them being distinct, the proposed algorithm takes exceptionally less time and space for executing than the other compared algorithms.

References

1. Yang, Y., Yu, P., Gan, Y.: Experimental Study on the Five Sort Algorithms. IEEE, New York (2011). ISBN: 978-1-4244-9439-2
2. Setiawan, R.: Comparing Sorting Algorithm Complexity Based on Control Flow Structure. IEEE, New York (2016). ISBN: 978-1-5090-3352-2

3. Kalid.: Sorting Algorithms, Better Explained. <https://betterexplained.com/articles/sorting-algorithms/>, Accessed 19 Aug 2018
4. Cormen, T H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to Algorithms, 3rd edn. MIT Press, Cambridge (2009). ISBN: 9780262259460
5. Edjlal, R., Edjlal, A., Moradi, T.: A Sort Implementation Comparing with Bubble Sort and Selection Sort. IEEE, New York (2011). ISBN:978-1-61284-840

A Survey on Wireless Optical Communication: Potential and Challenges



Sanjeev Kumar and Preeti Singh

Abstract With the increase in the number of users, the demand for bandwidth is growing in mobile communication. The RF-based technologies have limitations such as a lower data rate, security issues, an expensive licensing, a congested spectrum and a high cost of installation. To overcome these issues, wireless optical communication is an alternative. This paper focuses on FSO (Free Space Optics) a part of WOC (Wireless Optical Communication). There are a number of advantages of wireless optical communication: huge modulation bandwidth, no spectrum licensing, high security, low cost and many more. In spite of these advantages, the performance of optical wireless communication is degraded due to certain effects in both FSO. This paper gives an inclusive survey of FSO communication systems.

Keywords Radio frequency · Free space optics · Visible light communication · Wireless optical communication

1 Introduction

Communication is simply a way of transmitting information. It can be classified into two types: wired communication and wireless communication. Wired communication is relatively less affected by adverse weather conditions when compared to wireless solutions. Wireless communication can be further divided into Radio Frequency (RF) and Wireless Optical Communication (WOC). WOC can be divided into Visible Light Communication (VLC) and Free Space Optics (FSO).

This paper focuses on wireless optical communication, i.e. FSO. In recent years, information and communication technologies have shown numerous growth and advancement. As the numbers of users are growing day by day, the demand for high-

S. Kumar · P. Singh (✉)
UIET, Panjab University, Chandigarh, India
e-mail: preeti_singh@pu.ac.in

S. Kumar
e-mail: sanjeev_esd@yahoo.in

Table 1 Comparison of wired and wireless communication

Property	Wired communication	Wireless communication	
		RF	FSO
Operating range	1200–1600 nm	3–300 GHz	800–1600 nm
Licensing	Required	Required	Not required
Electromagnetic (EM) interference	EM interference	EM interference	No EM interference
Connection	All direction	LOS and NLOS	Only LOS
Operating range	Short to long	Short to long	Short (indoor) to long
Secured network	Yes	Less secured	Yes
Sources of noise	Reflection	Appliances (electrical & electronics)	Light sources (sun light & ambient light)
Power requirement	Medium	Medium	Low
Coverage	Good	Good	Good

speed Internet and other services also increases. Due to this, capacity and bandwidth requirements are also increasing rapidly, and this creates congestion in the present RF spectrum. So, there is a requirement of shifting to optical carrier from the existing RF carrier. Therefore, WOC is the technology that transfers information from one point to another using optical carrier through an unguided channel, and channel may be free space or an atmosphere. WOC can be used for high-speed broadband connection, because it provides ease of implementation, abridged power consumption (half of RF), extremely high bandwidth, improved channel security and reduced size (10% of the diameter of RF antenna) [1]. WOC transmission principle is similar to optical fibre communication but the difference is that in WOC unguided channel is used to transmit data and guided optical fibre in the previous one. In 1876, the first wireless telephone system was demonstrated by Alexander Graham Bell [2]. Recent advancements in optoelectronics components boosted the FSO communication system and researchers have also shown great interest in solving existing spectrum congestion and last mile bottleneck problems. Many companies have been set up to provide high speed wireless optical devices. The comparative analysis of wired and wireless communication (i.e. RF and Optical) in terms of various performance parameters is given in Table 1.

The rest of the paper is organized as follows: Sect. 2 discusses the literature survey of FSO. It includes challenges faced by FSO communication, techniques to overcome those challenges and future scope of FSO.

2 FSO Communication

Free space optics can be used for both indoor and outdoor communication. Laser diodes are used as transmitter to transmit data using optical carrier, and photodiodes are used to receive those optical signals. For receiver, photodiode can be avalanche photodiode or PIN diode [3].

In FSO system block diagram as shown in Fig. 1, the optical source is used at the transmitter side. The signal travels through free space and is affected by atmospheric turbulence. After demodulation data is retrieved. Apart from several advantages, the performance of FSO is degraded by various losses; their causes, effects and proposed solutions for the same are tabulated in Table 2.

There Are Several Techniques to Overcome Challenges Faced by FSO Communication:

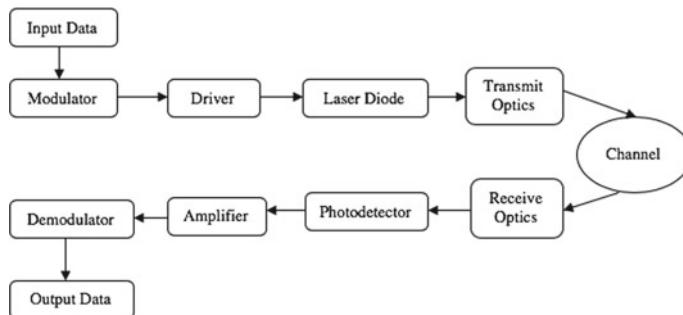


Fig. 1 Block diagram of FSO system

Table 2 Losses in FSO communication system

Type of loss	Cause	Effect	Solution
Absorption and scattering [23,24]	Rain, snow and fog	Attenuation (0.01–350) dB/km	Operating wavelength
Atmospheric turbulence [25,26]	Variation of temperature and pressure	Fluctuation in intensity and phase of received signal	Adaptive optics
Beam divergence loss [27]	Diffraction near receiver aperture	Signal loss	Efficient modulation
Misalignment [28]	Thermal expansion or vibration	Link failure	Aperture averaging
Background Noise [29]	Atmosphere	Power loss	Improved receiver
Pointing loss [30,31]	Satellite vibration or jitter	Link failure	Hybrid RF/FSO

Aperture Averaging: By increasing the size of receiver aperture, the influence of atmospheric turbulence can be abridged [4, 5]. As the receiver aperture is increased, the bit error rate reduces but due to increase in receiver aperture, the receiver noise also increases. Therefore, the diameter of the receiver aperture should be carefully chosen to balance both the effects.

Adaptive Optics: To reduce the atmospheric turbulence effect, adaptive optics can be used and output is an undistorted beam. Adaptive optics system consists of a closed-loop control system where conjugate of the atmospheric turbulence is used to correct the beam before transmitting into the atmosphere [6, 7]. Under very strong turbulent conditions, conventional adaptive optics approaches become less significant [8]. Received signal-to-noise ratio optimization is suitable for very strong turbulent conditions [9].

Diversity: The method to improve the reliability of a message signal using multiple numbers of communication channels with dissimilar characteristics is called diversity. Diversity can be classified into three types according to operation, i.e. time, space and frequency. An array is formed using small apertures instead of a single large aperture. This will improve bit error rate and link availability of the system [10]. M-ary Pulse Position Modulation (M-PPM) used with optical Multiple Input Multiple Output (MIMO) along with Low-Density Parity Check (LDPC) has shown improvement in coding gain of 57.8 dB [11]. Cooperative diversity immerged out to be an efficient technique [12, 13].

Modulation and Coding: Bandwidth efficiency and optical power efficiency are two main criteria that decide which scheme of modulation and coding should be used. Single carrier modulation is also known as binary modulation which is most widely used in FSO communication. Some common examples of binary modulation schemes are On-Off Keying (OOK) and Pulse Position Modulation (PPM). Adaptive threshold is used in OOK for good results in turbulent conditions [14]. Power efficiency can be improved by using M-PPM scheme [15]. Quadrature Amplitude Modulation (QAM) and Pulse Amplitude Modulation (PAM) are multi-carrier modulation techniques used in FSO communication [16]. Some other modulation techniques used in FSO communication are Subcarrier Intensity Modulation (SIM), Binary Phase Shift Keying (BPSK-SIM) [17], Differential Phase Shift Keying (DPSK) [18] and Orthogonal Frequency Division Multiplexing (OFDM) [19].

Hybrid RF/FSO: FSO performance is degraded by adverse weather condition and atmospheric turbulence. To overcome this problem, FSO can be used with RF to maintain link feasibility in all weather and any atmospheric condition, and this technique is called hybrid RF/FSO. So, hybrid RF/FSO is a good solution for adverse weather conditions [20]. Coding scheme using adaptive joint symbol rate is used in hybrid RF/FSO [21]. Hybrid RF/FSO link capacity and throughput are given in [22].

3 Conclusion

FSO communication has a good future because of its importance in mobile network backhaul, fibre backup, last mile access and many more. In the last decade, rapid growth has been observed in FSO communication in the global market due to fast deployment, quick availability and license-free operation. FSO communication can be used effectively in distance ranging from few centimetres to kilometres. FSO communication also finds application in chip-to-chip communication using FSOI (Free Space Optics Interconnects). Till date, more work is done on physical layer of FSO, so researchers should also work on upper layers of FSO like link layer, transport layer and application like TCP model. This technology definitely will bring revolution in the communication sector.

References

1. Chan, V.W.: Free-space optical communications. *J. Lightwave Technol.* **24**(12), 4750–4762 (2006)
2. Hochfelder, D.: Alexander graham bell. *Encylopedia Britannica* (2015)
3. Kaushal, H., and Georges K.: Free space optical communication: challenges and mitigation techniques (2015). [arXiv:1506.04836](https://arxiv.org/abs/1506.04836)
4. Perlot, N., Daniel, F.: Aperture averaging: theory and measurements. *Free-Space Laser Communication Technologies XVI*. Int. Soc. Opt. Photon. **5338**, 233–243 (2004)
5. Yura, H.T., and McKinley, W.G.: Aperture averaging of scintillation for space-to-ground optical communication applications. *Appl. Optics* **22**(11), 1608–1609 (1983)
6. Zocchi, F.E.: A simple analytical model of adaptive optics for direct detection free-space optical communication. *Opt. Commun.* **248**(4–6), 359–374 (2005)
7. Barbier, P.R., David, W.R., Mark, L.P., Penelope, P.D.: Performance improvement of a laser communication link incorporating adaptive optics. In *artificial turbulence for imaging and wave propagation*. Int. Soc. Opt. Photon. **3432**, 93–103 (1998)
8. Baranova, N.B., Mamaev A.V., Pilipetsky N.F., Shkunov V.V., Zel'dovich B.Y.: Wave-front dislocations: topological limitations for adaptive systems with phase conjugation. *JOSA* **73**(5), 525–528 (1983)
9. Vorontsov, M.A., Sivokon, V.P.: Stochastic parallel-gradient-descent technique for high-resolution wave-front phase-distortion correction. *JOSA A* **15**(10), 2745–2758 (1998)
10. Mohammad, N.S., Uysal M., Kavehrad M.: BER performance of free-space optical transmission with spatial diversity. *IEEE Trans. wirel. Commun.* **6**(8) (2007)
11. Djordjevic, I.B., Vasic, B., Neifeld, M.A.: Multilevel coding in free-space optical MIMO transmission with Q-ary PPM over the atmospheric turbulence channel. *IEEE Photon. Technol. Lett.* **18**(14), 1491–1493 (2006)
12. Karimi, M., Kenari, M.N.: BER analysis of cooperative systems in free-space optical networks. *J. Lightwave Technol.* **27**(24), 5639–5647 (2009)
13. Safari, M., Uysal, M.: Relay-assisted free-space optical communication. *IEEE Trans. Wirel. Commun.* **7**(12), 5441–5449 (2008)
14. Brown, W.C.: Optimum thresholds for optical On-Off keying receivers operating in the turbulent atmosphere. In *free-space laser communication technologies IX*. Int. Soc. Opt. Photon. **2990**, 254–262 (1997)
15. Yuen, J.H.: Autonomous Software-Defined Radio Receivers for Deep Space Applications, vol 13. Wiley, Amsterdam (2006)

16. Hranilovic, S., David, A.J.: A multilevel modulation scheme for high-speed wireless infrared communications. Proc. IEEE Int. Symp. Circuits Syst. **6**, 338–341 (1999)
17. Faridzadeh, M., Gholami, A., Ghassemlooy, Z., Rajbhandari, S.: Hybrid PPM-BPSK subcarrier intensity modulation for free space optical communications. In: 16th European Conference on Networks and Optical Communications (NOC), pp. 36–39. IEEE (2011)
18. Sinsky, J.H., Adamiecki, A., Gnauck, A., Burrus, C.A., Leuthold, J., Wohlgemuth, O., Chandrasekhar, S., Umbach, A.: RZ-DPSK transmission using a 42.7-Gb/s integrated balanced optical front end with record sensitivity. *J. Lightwave Technol.* **22**(1), 180 (2004)
19. Djordjevic, I.B., Vasic, B., Neifeld, M.A.: LDPC coded OFDM over the atmospheric turbulence channel. *Opt. Express* **15**(10), 6336–6350 (2007)
20. Kim, I.I., Korevaar, E.J.: Availability of free-space optics (FSO) and hybrid FSO/RF systems. In optical wireless communications IV. *Int. Soc. Opt. Photon.* **4530**, 84–96 (2001)
21. Kashyap, A., and Shayman, M.: Routing and traffic engineering in hybrid RF/FSO networks.” In Communications, ICC 2005. IEEE International Conference on, vol. 5, IEEE, (2005) 3427–3433
22. Liu, B., Liu, Z., Towsley, D.: On the capacity of hybrid wireless networks. In: INFOCOM Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies, vol. 2, pp. 1543–1552. IEEE (2003)
23. Weichel, H.: Laser beam propagation in the atmosphere, vol. 3. SPIE Press (1990)
24. Kim, I.I., Achour, M.: Free-space links address the last-mile problem. *Laser Focus World* **37**(6), 121–130 (2001)
25. Andrews, L.C., Phillips, R.L.: *Laser Beam Propagation Through Random Media*, vol. 152. SPIE Press, Bellingham (2005)
26. Rao, R.Z.: Scintillation index of optical wave propagating in turbulent atmosphere. *Chin. Phys. B* **18**(2), 581 (2009)
27. Kaushal, H., Kumar, V., Dutta, A., Aennam, H., Jain, V.K., Kar, S., Joseph, J.: Experimental study on beam wander under varying atmospheric turbulence conditions. *IEEE Photon. Technol. Lett.* **23**(22), 1691–1693 (2011)
28. Liu, X.: Free-space optics optimization models for building sway and atmospheric interference using variable wavelength. *IEEE Trans. Commun.* **57**(2), 492–498 (2009)
29. Katz, J.: Planets as background noise sources in free space optical communications. (1986)
30. Sannibale, V., Gerardo, G.O., William, H.F.: A sub-hertz vibration isolation platform for a deep space optical communication transceiver. Free-space Laser communication technologies XXI, International Society for Optics and Photonics **7199**, 71990K (2009)
31. Andrews, L.C.: Aperture-averaging factor for optical scintillations of plane and spherical waves in the atmosphere. *JOSA A* **9**(4), 597–600 (1992)

Breast Cancer Detection with a Ring-Shaped DRA (Dielectric Resonator Antenna)



Gagandeep Kaur and Amanpreet Kaur

Abstract This paper presents simulation results of tumor detection by placing a phantom parallel to square ring-shaped DRA. This is based on the difference between malignant and normal breast tissues electrical properties. For this purpose, a square ring-shaped DRA with plus DGS in ground is designed. This has volumetric dimensions of $60 \times 60 \times 5.67 \text{ mm}^3$. A spherical breast phantom model is designed with varying dielectric properties. The size of phantom is 44 mm with skin layer of 2 mm followed by fatty layer of 40 mm, followed by 6 mm diameter tumor toward the center of the phantom. The designed DRA covers bands of 4.6–5 GHz and 5.6–11.9 GHz and impedance bandwidths of 400 MHz and 6.3 GHz, respectively. The simulation results excited in front of the breast phantom show significant variations in terms of S11 in presence and absence of tumor. This allows the detection of tumor in phantom. The SAR results show average SAR_10 W/Kg, making antenna safe for biomedical applications.

Keywords DRA · DGS (defected ground structure) · Aperture coupled · Spherical breast cancer · Ultra-wideband · SAR · CST MWS 2017

1 Introduction

Breast cancer is the main leading cause of death among women. Breast cancer occurs due to unwanted growth of cells in the form of lumps (i.e. cancerous cells) around the normal cells. Due to these cancerous cells, tumors can be felt in the form of swelling lump in the breast density or can be observed with the help of various breast cancer detection methods with different tools but commonly used methods are X-ray mammography, ultrasound, and magnetic resonance imaging [1, 2]. The existing breast cancer detection techniques are not beneficial due to poor sensitivity and accuracy level. To overcome these types of precincts, microwaves are used in

G. Kaur (✉) · A. Kaur
ECED, TIET Patiala, Punjab, India
e-mail: gagandhindsa2127@gmail.com

UWB frequency area. This method is based on the dielectric properties variations of normal and tumors tissue. It is because there is more water content in malignant tissue than in normal tissue. When the spherical breast phantom is illuminated by an UWB pulse, more scattering takes place in the presence of high water content tumorous cell [3, 4]. This paper presents a heterogeneous breast model with three layers that contain 2 mm skin layer followed by fat and then tumor of 6 mm diameter. The spherical breast phantom is composed of three different layers namely skin, fatty tissue and tumor. Skin is the uppermost layer with dielectric constant of 36.58, conductivity of 2.34S/m and mass density of 1109 kg/m³. Fatty tissue is placed just behind the skin layer with dielectric constant of 4.83, conductivity of 0.262S/m, mass density of 911 kg/m³. The tumor cell is embedded inside the fatty tissue with dielectric constant of 67, conductivity of 49S/m and mass density of 84kg/m³ [5, 6]. This phantom is placed parallel to the proposed ring-shaped DRA operating at a UWB. In order to check the safeness of the proposed set up for tumor detection in human breast, the specific absorption rate results over the breast are presented at 4 GHz and 8 GHz, respectively, and the SAR results are found to be under the safe limits of human exposure, i.e., less than 16 W/kg. The detection model compares the reflected power recorded from close proximity area of the breast phantom in two cases: one with tumor and other without tumor. The antenna and phantom designing and all the relevant simulations are done using CST MWS V'17.

1.1 Proposed Antenna Structure with Simulated Results

The volumetric dimensions of simulated antenna design are $60 \times 60 \times 5.67$ mm³. The ultra-wideband antenna is composed of an aperture coupled feed loaded with hollow square-shaped resonator of alumina with permittivity of 9.8, thickness of 4 mm, and loss tangent of 0.0002 mounted on an FR4 substrate with permittivity of 4.4, wideness of 1.57, and loss tangent of 0.025 as shown in Fig. 1a. The hollow-shaped DRA used to increase the impedance bandwidth of antenna. The ground structure is printed above the dielectric substrate as shown in Fig. 1b. A DGS is realized on the ground and is used to interrupt the current distribution which depends on the outline of defect. So a plus (+) shaped DGS is etched from the center of the ground plane as shown in Fig. 1b, which is able to decrease the size and cross-polarization of the antenna. The best resonance is achieved when DRA is perfectly centered on the top of the aperture slot. The aperture coupled feed line is used for the excitation of antenna as shown in Fig. 1c. The guided wave propagating along the transmission line is coupled through the slot at resonant modes of the DRA. The feed line's width affects the coupling of slot, and thin feed line is preferred for strong coupling, i.e., 4 mm and obtained better results in terms of bandwidth, able to get an ultra-wideband with frequency of 4.6–5 GHz and 5.6–11.9 GHz with impedance bandwidth 400 MHz and 6.3 GHz, respectively. In Fig. 1a, c, dielectric material is shown in gray and copper material in white color. The whole measurements of labeled figures are cited in Table 1.

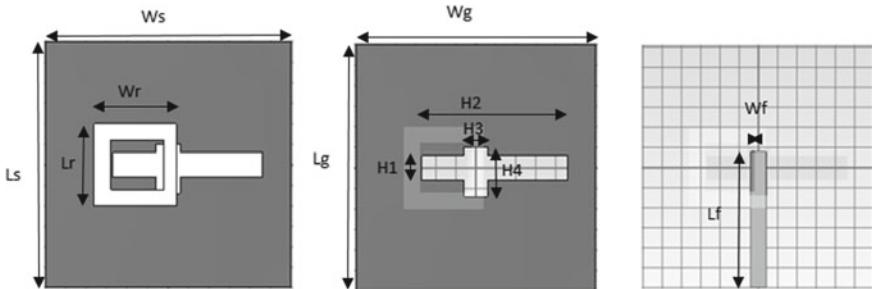


Fig. 1 DRA antenna geometry: **a** substrate with DRA ring, **b** ground plane with aperture, and **c** feed line

Table 1 Antenna's parameters

Parameter	Values (mm)
Substrate ($L_s \times W_s$)	60x60
Ground ($L_g \times W_g$)	60x60
feed-line ($L_f \times W_f$)	26x4
H_1	6
H_2	36.5
H_3	6

2 Simulated Results

The proposed DRA was designed and simulated in CST V'2017. The simulation results are achieved by optimization of ground structure, DGS, and DRA ring, and then antenna was matched to 50 ohms impedance. The outcomes are specified in the form of S_{11} , gain, VSWR, SAR, etc.

2.1 Return Loss

Figure 2 illustrates the simulated plot of return loss (dB) of the antenna structure. The graph is plotted between S_{11} parameter w.r.t. frequency (GHz). The antenna covers a band from 4.6 to 5.0 GHz and 5.6–11.9 GHz with impedance bandwidth 400 MHz and 6.3 GHz, respectively. The antenna shows resonance at 4.8 GHz and 7.68 GHz with peak return losses of -25.8 dB and -31.38 dB, respectively.

2.2 VSWR

Voltage standing wave ratio conveys information about the level of impedance matching between the transmission line. For better impedance matching, lesser VSWR is needed so more power is supplied to antenna. VSWR is calculated for efficient transmission in-between 1 and 2 for the first frequency band and 3 to 4 for the second frequency band as marked in the Fig. 3. The VSWR is 1.43, 1.18, 1.04, and 1.6 at resonance frequencies of 4.8 GHz, 6.03 GHz, 7.9 GHz, and 10.07 GHz, respectively.

2.3 Antenna Gain

The gain of antenna indicates that how well a signal can send or receive in a specified direction by the antenna. For the radiation of the antenna in a particular direction generally requires high gain (greater than 3dB) which is easily achieved by the proposed antenna design. Figure 4 shows the peak gain of 4.582 dB at the resonant frequency of 8 GHz. Figure 5 indicates the antenna's broadband gain over the entire UWB that it radiates.

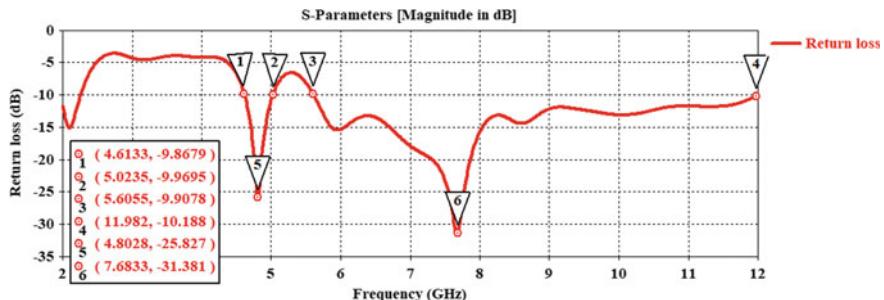


Fig. 2 Simulated return loss

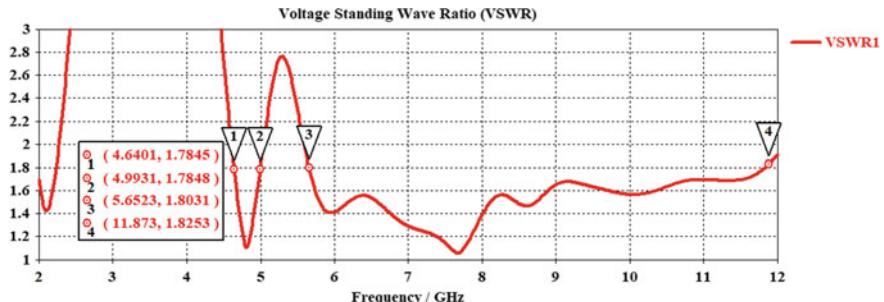


Fig. 3 VSWR of DRA

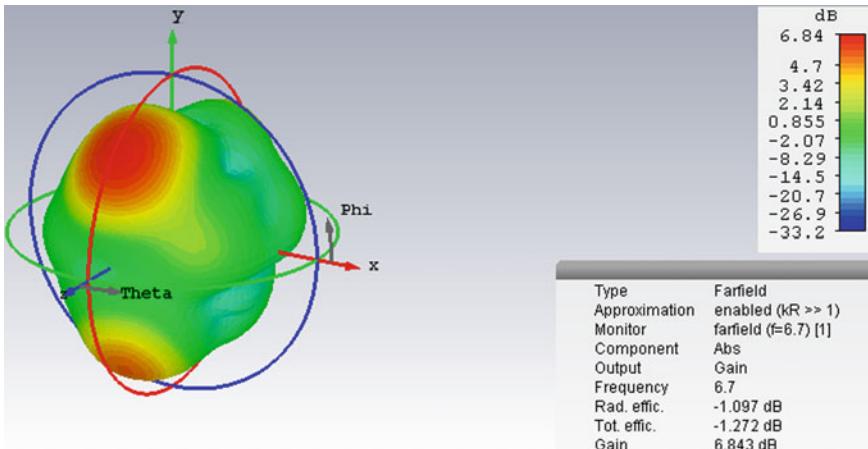


Fig. 4 Peak gain at 6.7 GHz frequency

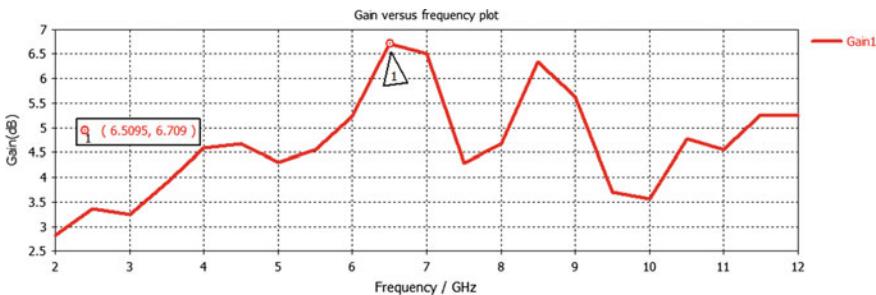


Fig. 5 Broadband gain

2.4 Design of the Breast Phantom and Simulations for Tumor Detection

The principal of microwave imaging for detection of breast tumor depends upon the dielectric contrast of the electrical properties difference between malignant and normal breast tissue. Initially, a spherical breast phantom model is designed. This three-layered spherical breast phantom model consists of a skin layer with 2 mm thickness and two different layers, i.e., fatty layer and tumor of diameter 40 mm and 6 mm, respectively, inserted inside the skin tissue. Figure 6a represents the side view of DR antenna, and Fig. 6b represents the heterogeneous layers of breast phantom. A spherical tumor is placed at location of (0, 0 and 20) inside the skin layer as shown in Fig. 6b. The dielectric contrast between normal skin and tumor is normally in 5:1 because malignant tissues in breast exhibit large water content (high permittivity) as compared to normal one and easily detect the location of tumor. The first most layer

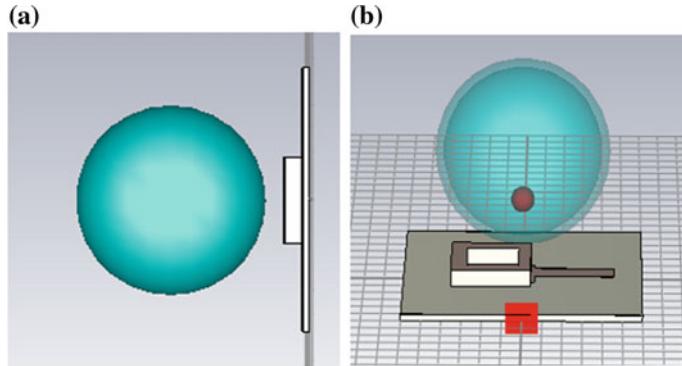


Fig. 6 **a** Side view. **b** Front view of proposed DRA parallel to phantom

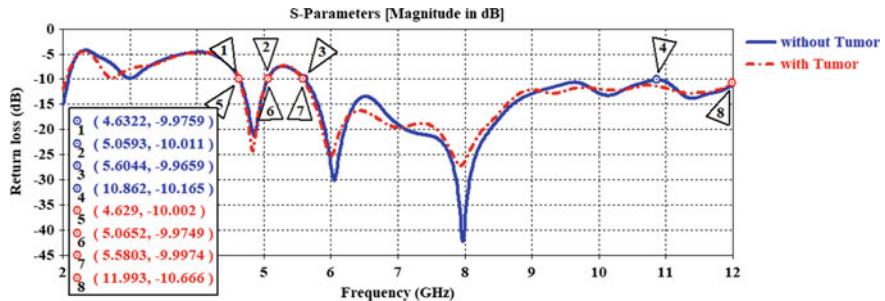


Fig. 7 Comparison plot of DRA with tumor and without tumor

of skin has an epsilon 36.587, electric conductivity of 2.34 S/m, and mass density of 1109 kg/m³. The second (middle) layer is fatty tissue has an epsilon of 4.8393, electric conductivity of 0.262 S/m, and mass density of 911 kg/m³; and the third layer of tumor tissue has an epsilon of 67, electrical conductivity of 49 S/m, and mass density of 84 kg/m³. The proposed antenna structure is placed parallel to the spherical breast phantom along Z axis and the distance between antenna and phantom is 10mm. The UWB gaussian pulses are transmitted by the proposed antenna toward the breast phantom and recorded by the same antenna. Difference in the return loss characteristics is observed with tumor and without tumor due to high scattering in presence of tumor.

Figure 7 shows a comparative plot of S11 parameters against frequency for both the cases of with and without tumor. It is observed that high water content malignant tissue shows more scattering as compared to the normal breast tissue. The proposed antenna design shows more variation at resonant frequencies of 4.85 GHz, 6.02 GHz and 8 GHz for the cases with and without tumor, show more reflection and detection of 8 mm (diameter) tumor is easily recognizable due to large variation in magnitude (in terms of return losses) as mentioned in Table 2.

2.5 Specific Absorption Rate (SAR)

The SAR is an amount of the power absorbed by the medium per unit of the mass. In breast cancer detection, SAR is the standard that can put some limit to the maximum amount of power absorbed by the breast. SAR can be considered by Eq. (1):

$$SAR = \frac{|E|^2 \sigma}{\rho}, \quad (1)$$

where σ (S/m) is material conductivity, E (V/m) is electric field intensity and ρ (Kg/m³) is mass density. From the equation, it is clear that the focusing point is the amplitude of E and not phase. Figure 7 shows the normalized SAR distribution with a single DRA through phantom at 4 GHz frequency is 11.9 W/kg (as shown in Fig. 8a and at 8 GHz frequency is 9.034 W/kg (shown in Fig. 8b). The SAR calculations at different coordinate points have been mentioned in Tables 3 and 4 at 4 GHz and 8 GHz frequencies, respectively.

Table 2 Return loss deviation

Frequency(GHz)	Return loss in absence of tumor (dB)	Return loss in presence of tumor (dB)	Return loss deviation (dB)	Tumor detection
4.8	-21.2	-24.3	3.1	Detectable
6.03	-29.5	-24.6	4.9	Detectable
7.9	-41.05	-26.93	14.12	Detectable
10.07	-13.2	-12.09	1.11	Detectable

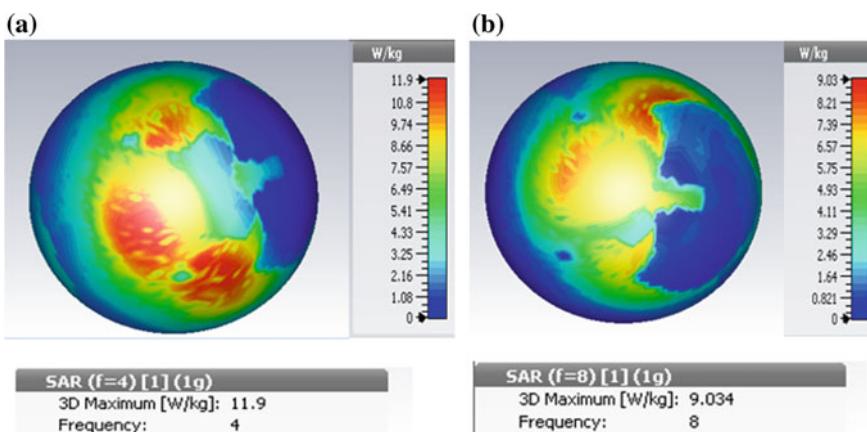


Fig. 8 SAR results **a** at 4 GHz frequency and **b** at 8 GHz frequency

Table 3 Simulated values of SAR results at 4 GHz at different positions in the breast

X-position	Y-position	Z-position	SAR(W/Kg)
4.62	9.058	1.05	4.05
8.86	3.35	1.18	7.62
1.34	1.54	1.46	1.13
1.95	6.52	2.19	8.15
2.18	9.47	2.94	2.08

Table 4 Simulated values of SAR results at 8 GHz at different positions in the breast

X-position	Y-position	Z-position	SAR(W/Kg)
4.87	2.33	1.07	6.80
5.22	6.64	1.17	6.41
4.95	1.48	1.65	6.94
4.90	1.88	2.19	3.53
3.69	2.13	2.84	2.09

3 Conclusion

This paper presents a ring-shaped DRA loaded with aperture coupled feed line. For the enhancement of impedance bandwidth, a plus-shaped slot is carved out from the ground surface. The proposed DR antenna works on the principle of microwave imaging for detection of breast detection. In this method, with the help of the dielectric difference of electrical properties, the difference between malignant and normal tissue of the breast is used to detect the position of the tumor. The designed antenna covers an ultra-wideband of frequency from 4.6 to 5 GHz and 5.6–11.9 GHz with impedance bandwidths of 400 MHz and 6.3 GHz, respectively. This antenna is used for microwave imaging applications, by designing a breast phantom with a tumor and electrical properties similar to that of a human breast. Then the designed antenna is used as a trans-receiver to transmit UWB pulses to the phantom and receive backscattered signals from it. An appreciable variation in the S11(dB) values of the two cases of with and without tumor leads to the conclusion that illuminating a human body part with the proposed antenna emitting UWB frequencies in the microwave region can be used to detect tumors in the human body. The antenna proposed in this section is safe for the biomedical purpose as it has an SAR below 16 W/kg.

Acknowledgements This work would not have been possible without the support of Thapar Institute of Engineering and Technology (TIET), Patiala.

References

1. Zhang, Z., Liu., Q.: Three-dimensional nonlinear image reconstruction for microwave biomedical imaging. *IEEE Trans. Biomed. Eng.* **51**, 544–548 (2004)
2. Ali, I., Wani, W.A., Saleem, K.: Cancer Scenario in India with Future Perspectives. *Cancer Therapy.* **8**(8), 56–70 (2011)
3. Karli,R., Ammor, H: Miniaturized UWB microstrip patch antenna with t-slot for detecting malignant tumors by microwave imaging. *Int. J. Microw. Opt. Technol.* **9** (3), 214–220(2014)
4. Palantel, E.: Early stage cancer detection technique considering the reflected power from breast tissues. *ARPEN J. Eng. Appl. Sci.* **10**(17), 7361–7367(2015)
5. Xu, M.-H., Wang, L.-H.V.: Time-domain Reconstruction for Thermo-acoustic Tomography in a Spherical Geometry. *IEEE Trans. Med. Imag.* **21**, 814–822 (2002)
6. Sharma, S.K., Brar, M.K.: Aperture coupled pentagon shaped dielectric resonator antennas providing multiband and wideband performance. *Microw. Opt. Technol. Lett.* **55**, 395–400 (2013)

Simulation and Analysis of Distinct Apodized Profiles Using Fiber Bragg Grating for Dispersion Compensation at 100 Gbps Over 120 km



Ashwani Sharma, Inder Singh, Shalini Sharma, Robin Thakur and Raj Kumar Saini

Abstract Due to the advancements in modern technology, there is an immense use of Internet worldwide. For the smooth transmission of the information from one end to other, systems having characteristics of more bandwidth and high capacity are required. Hence, optical fiber is used for this purpose. Like all other transmission media, optical fiber also suffers from signal deteriorating factor that includes attenuation, dispersion, and other nonlinear effects. Hence, there is need for their compensation that includes utilizing dispersion compensation techniques like DCF, FBG, EDC, and digital filters. The work in this paper evaluates the performance of distinct apodized profiles for FBG. Basically, three types of profiles: Gaussian, uniform, and hyperbolic tangent are executed and investigated at a data rate of 100 Gbps over a transmission distance of 120 km. After analyzing the outcomes, these techniques are then compared in order to get the best apodization scheme for FBG to implement so as to achieve less dispersion at the receiver end.

Keywords OptiSystem · Dispersion · FBG · IDCFBG · Chromatic dispersion · Bit error rate · Q-factor · Gaussian · Uniform · Hyperbolic tangent

1 Introduction

FBG is turned out to be the predominant segment in the optical communication systems. FBGs are considered as filters that are based on Bragg's reflection theory. Their basic principle is to reflect the light having shorter wavelengths and transmits the other leftover wavelengths. During the propagation of light inside the fiber Bragg

A. Sharma (✉) · S. Sharma
School of Electrical and Computer Sciences, Shoolini University, Solan, Himachal Pradesh, India
e-mail: asharma7772001@gmail.com

A. Sharma · I. Singh
School of Computer Science Engineering, UPES, Dehradun, UK

R. Thakur · R. K. Saini
School of Mechanical and Civil Engineering, Shoolini University, Solan, Himachal Pradesh, India

grating, overall reflection occurs at Bragg's wavelength while not affecting the left-over wavelengths. A large portion of few side lobes existing in reflection spectrum is also reflected by the fiber Bragg grating. Hence, to ensure the better transmission of information between transmitter and receiver ends, suppression of these side lobes is necessary [1–4]. Elimination of crosstalk between information carrying channels also requires the repression of these side lobes in order to use it in applications like wavelength division multiplexing. Apodization is a technique normally used to suppress side lobes in the spectrum by moderately altering the coupling coefficient's amplitude along the length of the grating. For distinct fiber Bragg grating applications, distinct parameters are required, thereby causing a change in the profiles of apodized function. A variety of apodized profiles are explored for compensation of dispersion in fiber optics. Most common apodization functions available are raised cosine, Gaussian, sine, etc. All of these techniques have their unique features and distinct methods of fabrication. Except for the compensation of dispersion in fiber optics, FBG also has applications in sensor technology.

The length of the grating and the grating strength are the two main properties of FBG that need to be controlled. These are controlled by controlling the values of bandwidth, reflectivity, and strength of side lobes. Apodization has a significant role in suppressing the side lobes and narrowing the bandwidth. The spectral properties of FBG can be explained with the help of coupled mode theory. Another technique known as transfer matrix method divides the whole grating into subsections, and their inputs and outputs are then calculated and multiplied to get the overall response of the grating [5–7].

As the light propagates inside the optical fiber, fiber Bragg grating reflects the narrow band spectral components which are at Bragg's wavelength. This wavelength is given by Eq. (1) shown below [8, 9]:

$$\lambda_B = 2n_{eff}\Lambda, \quad (1)$$

where n_{eff} = effective refractive index, and Λ = grating period.

The mutation in refractive index along with grating period can be normally expressed as given below in Eq. (2):

$$n(z) = n_0 + \Delta n(z) \cos\left(\frac{2\pi}{\lambda} z + \theta(z)\right), \quad (2)$$

where $\Delta n(z)$ = envelope of refractive index modulation, n_0 = refractive index of core, and $\theta(z)$ = phase of refractive index modulation.

In case of unchirped gratings, $\theta(z) = 0$.

2 Simulation Model

The need for high bandwidth and system with high capacity for a longer distance at high bit rate is always desirable. Thus, there is a requirement of updation of optical fiber to provide compensation for the signal limiting factors like losses and dispersion. To do so, distinct compensation methods are utilized like DCF, FBG, etc. Here we are using a dispersion compensation module, which is the combination of EDFA, FBG, and IDCFBG. Here, the FBG is executed in three different apodized profiles, i.e., Gaussian, uniform, and hyperbolic tangent. For the above referred profiles, the software naming OptiSystem 7.0 has been used.

The setup for the simulation incorporates transmitter, optical channel, and recipient. Bits are transmitted in zeros and ones form, and hence, for this pseudo-random bit sequence, generator is employed to generate bits at 100 Gbps rate. Further, a nonreturn to zero pulse generator is used for changing that binary data into electrical pulses. To travel over longer distances, these electrical pulses require modulation. Hence, for this modulation purpose, a carrier signal of high frequency is generated by using a continuous laser and modulated using MZ modulator. In order to mitigate the effects of dispersion, FBG is used for compensating dispersion. A single-mode fiber along with dispersion compensation module comprises the optical channel. While propagating, the signal undergoes losses, and to reduce this attenuation, EDFA is utilized. The recipient end requires electrical signal so the incoming signal is converted back to electrical form by using pin photodetector. Further for the isolation of carrier, a low-pass filter is used. Eventually, a bit error rate analyzer analyzes the signal. For different apodized profiles, these outcomes have been compared to get the best profile. Simulation setup for the described model is shown in Fig. 1.

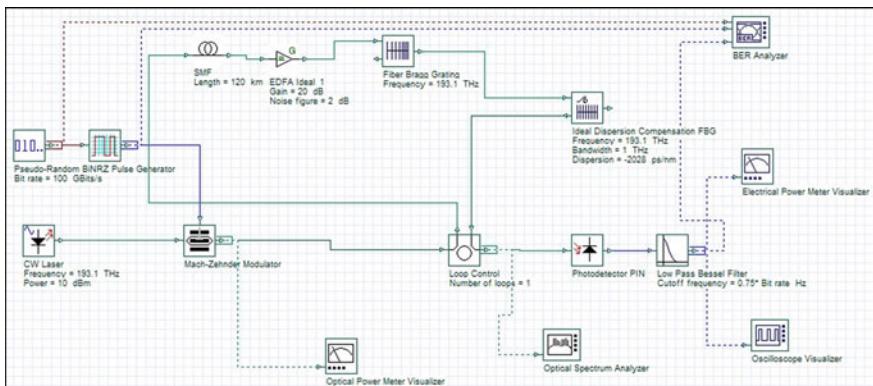


Fig. 1 Simulation setup for distinct apodization profile

Table 1 Parameters for simulation

Sr. no.	Parameter	Value
1	Frequency (THz)	193.1
2	Gain (dB)	20
3	Bit rate (Gbps)	100
4	Noise(dB)	2
5	Extinction ratio	30
6	Sample rate (THz)	6.4
7	Power (dBm)	1–10

Table 2 Parameters for IDCFBG

Sr. no.	Parameter	Value
1	Fiber length (in km)	120
2	Differential group delay (in ps/km)	3
3	Differential slope (in ps/2/)	0.008
4	Attenuation (in db/km)	0.2
5	Dispersion (in ps/nm/km)	17

Table 3 Simulation parameters of fiber Bragg grating

Sr. no.	Parameter	Value
1	Frequency (THz)	193.1
2	Apodization	Uniform, Gaussian, and tanh

Table 1 is showing the simulation parameters of single-mode fiber whereas simulation parameters for IDCFBG and FBG are tabulated in Tables 2 and 3, respectively.

3 Results and Discussions

Simulations are performed using OptiSystem 7.0. Initially, fiber Bragg grating is apodized with uniform profile. The whole setup is executed at 120 Gbps over a transmission distance of 100 km, and then, the outcomes of the uniform apodized profile have been analyzed. The performance has been observed in terms of the quality factor, eye height, and bit error rate.

The distinct outcomes of uniform apodized profile are displayed in the form of the tabular form shown in Table 4. Another apodization techniques used are Gaussian and hyperbolic tangent. By applying the respective mathematical function on the fiber Bragg grating and varying input power from 1 to 10 dBm, the outcomes have been observed in terms of the quality factor, eye height, and bit error rate. These results are displayed in the form of table in Tables 5 and 6, respectively.

Table 4 The outcomes of uniform apodization profile

Input power (dbm)	Q-factor	BER
1	5.96221	1.24E-09
2	6.45793	5.31E-11
3	6.93019	2.10E-12
4	7.49575	3.28E-14
5	8.00952	5.71E-16
6	8.57587	4.85E-18
7	9.1059	4.20E-20
8	9.55193	6.23E-22
9	9.76573	7.78E-23
10	9.65732	2.28E-22

Table 5 The outcomes of Gaussian apodization profile

Input power (dBm)	Q-factor	BER
1	5.75493	4.31E-09
2	6.19799	2.85E-10
3	6.64245	1.54E-111
4	7.1765	3.58E-13
5	7.66497	8.93E-15
6	8.22575	9.65E-17
7	8.8143	5.97E-19
8	9.40435	2.58E-21
9	9.9585	1.14E-23
10	10.2808	4.25E-25

Table 6 The outcomes of tanh apodization profile

Input power	Q-factor	BER
1	5.955	1.29E-09
2	6.45224	5.51E-11
3	6.92726	2.14E-12
4	7.4805	3.70E-14
5	7.99785	6.28E-16
6	8.57299	4.97E-18
7	9.11634	3.81E-20
8	9.58397	4.57E-22
9	9.82077	4.51E-23
10	9.72923	1.13E-22

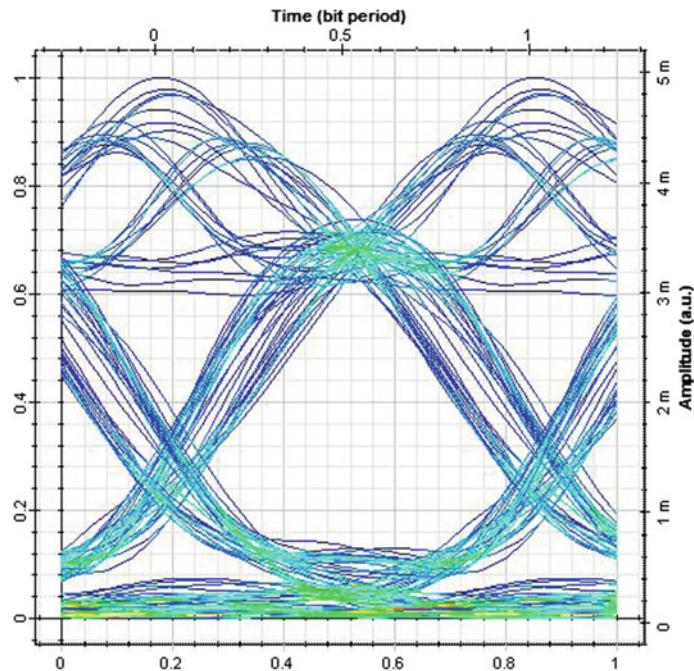


Fig. 2 Eye diagram of uniform function

Respective eye diagrams of uniform, Gaussian, and hyperbolic tangent apodized profiles are shown in Figs. 2, 3 and 4.

Distinct outcomes from distinct apodized functions are compared in terms of their respective BER and quality factor by varying input power from 1 to 10 dBm. Their comparison is shown by using graphs in Figs. 5 and 6, respectively. Comparison of quality factors of these three profiles is displayed graphically in Fig. 5 whereas bit error rate comparison is also shown graphically in Fig. 6. It has been shown that as the input power rises, respective quality factors of all the three apodized profiles grow at a constant rate. But when the outcomes of bit error rate are analyzed, they showed different responses for all the three apodized profiles. For all the three profiles, it fluctuates between low and high. When Q-factor is concerned, Gaussian function is exhibiting better performance than the remaining profiles but when the bit error rate is considered to be the major factor, then hyperbolic tangent is the dominant apodized profile among them. After analyzing the graphs in terms of both quality factor and bit error rate, it is found that hyperbolic tangent apodized fiber Bragg grating has superior performance among others.

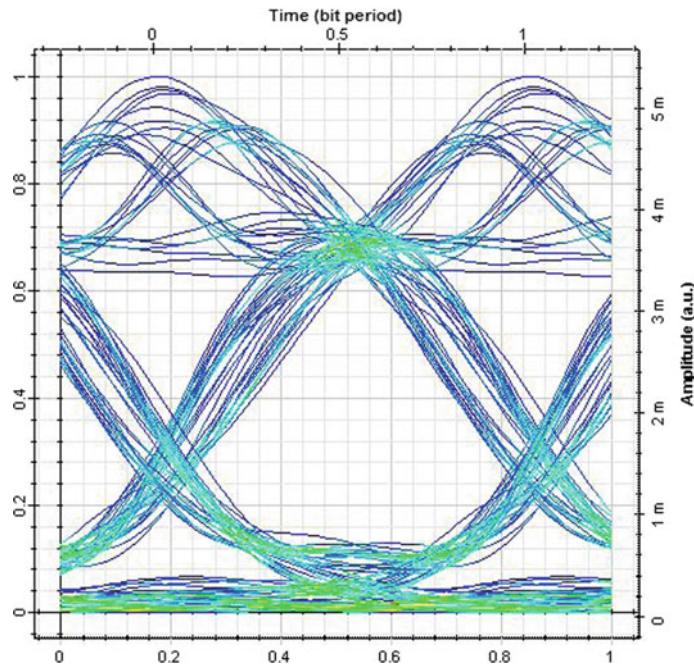


Fig. 3 Eye diagram of Gaussian function

4 Conclusion

The work in this paper completely revolves around executing the various available apodized profiles of fiber Bragg grating. Fiber Bragg grating can be used in distinct configurations and exhibits better outcomes than the other techniques for dispersion compensation. Except for configurations, alteration in core refractive index of fiber Bragg grating results in a better model for compensation of dispersion. This paper compared the outcomes of three distinct apodized profiles: uniform, Gaussian, and hyperbolic tangent in terms of both BER and quality factor. When input power grows, the corresponding quality factor of all three profiles also increases and Gaussian apodized profile rises to have the maximum Q-factor. But when bit error rate is considered to be the result-determining factor, then tanh exhibits minimum bit error rate. There is a trade-off between both the factors. Hence, it has been concluded that depending upon the application, apodized fiber Bragg grating can be used. But overall, the performance of hyperbolic tangent is far better than Gaussian and uniform apodized fiber Bragg grating.

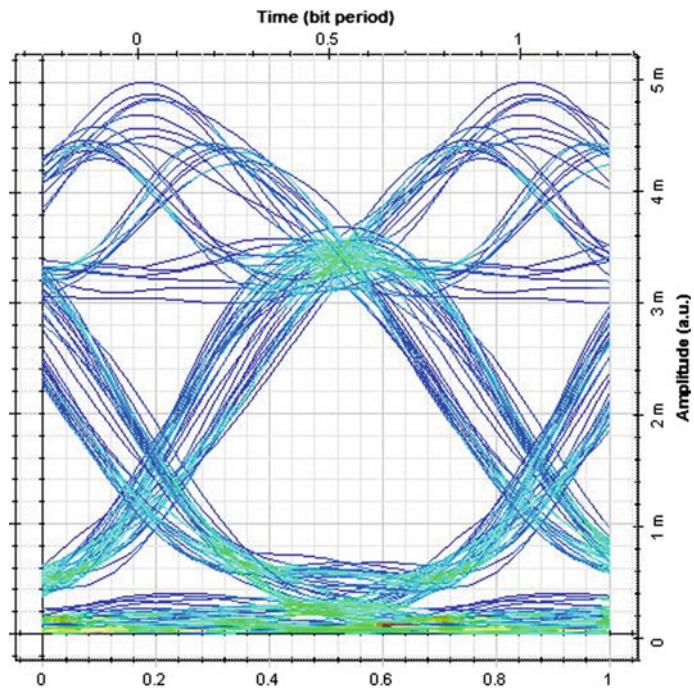


Fig. 4 Eye diagram of hyperbolic tangent (tanh) function

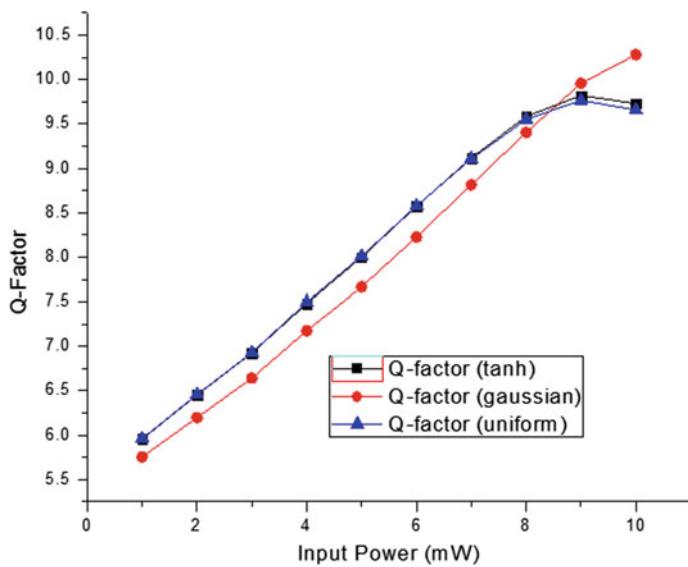


Fig. 5 Q-factor comparison

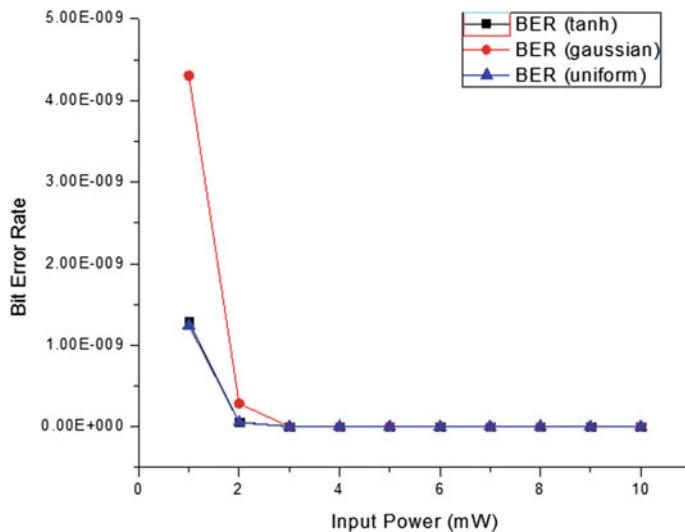


Fig. 6 Bit error rate comparison

References

- Sharma, A. et al.: Performance Analysis of Dispersion Compensation using Ideal Fiber Bragg Grating in a 100 Gbps single channel optical system. *Int. J. Eng. Sci. Res. (IJESR)* **07**(2), 421–430 (2018)
- James, F., Brennan, III: Broadband fiber Bragg gratings for dispersion management. *J. Opt. Fiber Commun. Rep.* 397–434 (2005)
- Zhu, L et al. (2001) System simulation of dispersion compensation with specially sampled fiber bragg grating. *Opt. Commun.* 89–93 (2001)
- Sumetsky, M. et al.: Fiber Bragg gratings for dispersion compensation in optical communication system. *J. Opt. Fiber Commun. Rep.* 256–278 (2005)
- Litchinitser, N.M et. al.: Fiber bragg gratings for dispersion compensation in transmission: theoretical model and design criteria for nearly ideal pulse recompression. *J. Lightw. Technol.* **15**(8), 1303–1313 (1997)
- Sun, N.H. et al.: Numerical analysis of apodized fiber bragggratings using coupled mode theory. *Prog. Electromag. Res. (PIER)* **99**, 289–306 (2009)
- Khalid, K.S. et al.: Simulation and analysis of Gaussian apodized fiber Bragg grating strain sensor. *J. Opt. Soc. Am.* 667–673 (2012)
- Singh, J. et al.: Design of Gaussian apodized fiber bragggrating and its applications. *Int. J. Eng. Sci. Technol.* **02**(05), 1419–1424 (2010)
- Prasad, G.S. et al.: A study on uniform and apodized fiber bragg gratings. *Int. J. Sci. Res. Dev. (IJSRD)* **03**(10), 855–857 (2015)

Development of Integrated Distance Authentication and Fingerprint Authorization Mechanism to Reduce Fraudulent Online Transaction



Vipin Khattri, Sandeep Kumar Nayak and Deepak Kumar Singh

Abstract From last numerous decades, there is a problem of counterfeit online transaction and poses as a foremost challenge of online transaction. Even though the concrete initiatives have been taken by researchers and governments, fraudsters acquire canny fashion to perform counterfeit online transaction. Fundamentally, fraudster steals the credentials of client and performs counterfeit online transaction. After scrutinizing, this research has stepped forward with a plan to trim down the counterfeit online transaction. This study propounds a working by blending two different workings which are authentication of distance and authorization of fingerprint to execute a genuine transaction. The leading key aspect behind the propound working is that after stealing credentials of the client, fraudsters could not perform counterfeit online transaction. In this study, to assess the impact, propound working is also theoretically implemented in some cases. After the assessment, it is found that the propound working is suitable to prevent counterfeit online transactions.

Keywords Counterfeit online transaction · Integrated distance and fingerprint · Authorization fingerprint · Authentication distance · Online transactions

V. Khattri · S. K. Nayak (✉)

Department of Computer Application, Integral University, Dasauli, Basha Kursi Road, Lucknow 226026, India

e-mail: nayak.kr.sandeep@gmail.com

V. Khattri

e-mail: vipinkhattri@gmail.com

D. K. Singh

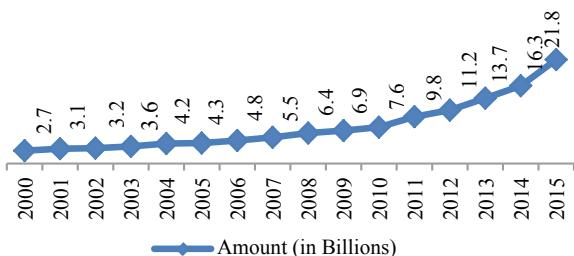
Jaipuria Institute of Management, Vineet Khand, Gomti Nagar, Lucknow 226010, India

e-mail: deepak.iiita@gmail.com

1 Introduction

The approach of online transactions is to digitally transfer the amount from source to target account, for example, online shopping. The online transaction approach is formulated with the objectives of automatizing the system, save time, workforce, paperwork, amplify security and comfort, and implementing a transparent system. Fraudsters steal money digitally or physically from online transaction. As a consequence, the tendency of fraudsters is to dig the flaws in the system and take benefits from it to perform counterfeit online transaction [1]. To conquer the issue of counterfeit online transaction, diverse methods are implemented. These methods are performed at two different security levels [2]. Primary level is authentication and secondary level is authorization. At the level of authentication, techniques of multi-factor authentication are implemented, e.g., visual cryptography [3], fingerprint [4], security token, one-time password (OTP) [5], and many more. At the level of authorization, real-time assessment of online transaction is implemented using numerous techniques such as Naïve Bayesian [6], artificial immune system [7, 8], genetic algorithm [9, 10], hidden Markov model [11, 12], and artificial neural network [13–16]. During investigating the statistics of counterfeit online transaction, every country is facing the problem of counterfeit online transaction. In India, around 37% of card holders are troubled due to counterfeit online transaction [17]. Figure 1 is showing the global loss of amount due to online transaction fraud. Loss of amount in 2014 was \$16.4 billion which was increased to \$21.8 billion in 2015, i.e., a 25% increase as compared to the previous year. The situation of loss of amount is happening after applying various security levels, i.e., authentication and authorization. To handle the situation by enhancing security levels, the authors propound a method by blending two different methods at primary level using multifactor authentication and at secondary level using fingerprint authorization during the online transaction. The authors split this study into five fragments. The second fragment shows the base of the propound work. The third fragment defines the propound method to diminish the counterfeit online transaction. The fourth fragment analyzes the outcome of the propound method when it is implemented on different case studies of counterfeit online transaction. The last fragment gives the concluded remark with future scope of this study (Fig. 1).

Fig. 1 Financial loss due to counterfeit online transaction [17] (Source 2016 Report-Nilson)



2 Background of Research

2.1 *Different Forms of Counterfeit Online Transaction*

Counterfeit Online Transaction of ATM. To perform counterfeit online transaction in ATM, fraudster plants the skimmer device in ATM at the top of the slot of card swap section. When the payment card is swapped by the user, all the credentials of payment card are copied to the skimmer device. The fraudster also plants a fake keypad on the actual keypad of the ATM machine. When the PIN number is entered by the user, the PIN number is copied to the false keypad. After completing the complete transaction, fraudster extracts the card details from the skimmer device to make a clone of card and extracts the PIN number from the false keypad. Using a cloned card with the PIN number, the fraudster performs the counterfeit online transaction at ATM or point of sale (POS) [18–20].

Counterfeit Online Transaction Using Stolen or Lost Payment Card. There are two situations in which fraudsters perform counterfeit online transaction either by getting lost payment card or by stolen payment card. In the first situation, the fraudster finds the lost card in which the details of the PIN number are also written. In the second situation, the fraudster watches the activities in which the user uses the card for payment and records the PIN number by either through hidden camera or using shoulder surfing. After recording the PIN number, the fraudster steals the payment card. This lost or stolen card is used to perform counterfeit online transaction [1, 20 21].

Counterfeit Online Transaction using Phone Fraud. In this category of fraud, the fraudster informs the user using the phone with false information such as your card has been blocked or due to security reason please verifies your card details. In this type of situation, the fraudster initiates the counterfeit online transaction at his end and through phone extracts all the credentials of payment card with OTP number and fill all the details of the transaction at his end and performs complete transaction [1, 20, 21]. Both the things that are extracting details from the user and counterfeit online transaction at fraudster ends are performed simultaneously or parallel or concurrently.

2.2 *Distance Authentication's Mechanism*

The author propounded a mechanism [2] to prevent counterfeit online transaction. With this concept, during the online transaction, the user must be physically present at the place where the transaction is being conducted. To ensure this, the author grabs the gap between the mobile phone of the user and the device on which transaction is performed. If this gap is within the critical gap (constraint by bank), then the user is permitted to precede transaction; otherwise, the transaction is denied during the initial authentication of the online transaction. The author concluded that when

the fraudster performs the counterfeit online transaction, it fails because of a large distance between user and fraudster.

2.3 Authorization of Online Transaction Using Fingerprint

To ensure the security and authorization for access, fingerprint authorization is a valid, secure, and unique technique. The same aspect can also be used during the online transaction for authorization. A most novel aspect of the fingerprint is that it is only one of its kind and cannot be simulated like signature/password. The authorization process ensures that both the impression should be the same. By utilizing the unique technique in the authorization of online transaction [4], the security, safety, trust, and convenience are augmented with reducing the risk of loss of money and credentials. It is utilized in important areas of business where security is the prime objective such as financial services, defense security, health services, etc. Fingerprint can be utilized as an authorization of online transaction to augment the security of online transaction system.

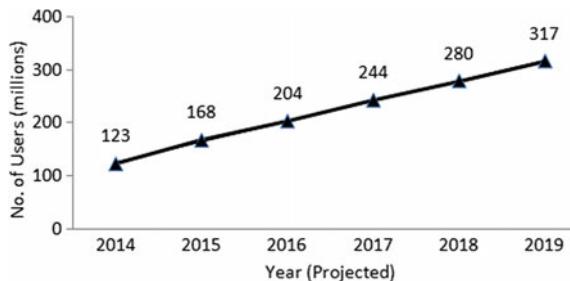
3 Propound Integrated Distance and Fingerprint Mechanism

The important encouragement of creating the propound mechanism is that the counterfeit online transaction could not begin or complete by fraudster even if the fraudster has all the stolen information of payment card. Another important encouragement of propound method is to enhance the security in terms of authentication and authorization and without acquaintance of user; fraudster could not initiate the counterfeit online transaction. The purpose of choosing the cell phone for authentication and authorization for online transaction is that one thing which is carried all the times by human is the mobile phone. It is an integral part of life. After analyzing the graph as given in Fig. 2, everyone can make a clear picture of increasing rate of use of mobile phone, and it is increasing with tremendous growth rate [22] (Fig. 2).

3.1 Elements of Propound Mechanism

Online Transaction Instruments. There are two transaction instruments used for implementing the propound mechanism. The first instrument is cell phone. This cell phone is used to open financial application for initiating the online transaction by the user using some code. Another use of cell phone in propound mechanism is that when the user initiates the transaction, the location of the cell phone is also sent to

Fig. 2 Users of cell phone [22]



the transaction server for authentication, and then, it authorizes the user by taking the fingerprint impression. The second instrument is the transaction instrument. Online transaction is performed using this instrument. This instrument can be of different types such as personal computer, laptop, point of sale, ATM, etc.

Financial Application. An application which is named as “Financial Application” is especially designed by financial institution for authentication and authorization of online transaction. The financial application must be installed on the cell phone. For authentication, the location of the cell phone is sent to the transaction server, and for authorization, the fingerprint impression which is taken from the user is sent to the transaction server. For security reasons, this financial application is linked with the registered mobile number. Another aspect of the security of the online transaction, no incoming/outgoing phone calls can be initiated during online transaction while using financial application.

Transaction Payment Card. Any online transaction whether it is performed using ATM or POS or e-commerce website requires a payment card. This payment card must be issued to the user by the financial institution to perform the online transaction.

Fingerprint Authorization. In the propound mechanism, validation of fingerprint impression during the online transaction for authorization is required. This validation is only achievable by a third party or financial institution. Therefore, the fingerprint impression should be recorded by the third party or financial institution and used for authorizing the online transaction.

Desirable Constraints of Propound Mechanism.

- Account of the user should be registered in the financial institution.
- User must have a payment card to perform online transaction issued by the financial institution.
- User should have a cell phone number and this number should be linked with an account in the financial institution.
- User should have a smart cell phone with the facility of fingerprint recognition or using additional device.
- Financial application should be installed in cell phone for performing online transaction.

- When the user initiated transaction, then it must be ensured that the gap between the user's cell phone and transaction instrument should be less than 2 km and the fingerprint impression should be valid.

3.2 Integrated Approach of Authentication and Authorization Propound Mechanism

The propound mechanism (as shown in the Fig. 3) acts a crucial role along with cell phone, gap between cell phone and transaction instrument, financial application, and fingerprint impression for reducing counterfeit online transaction. This method increases the defense for performing a secure online transaction. This is only possible when two aspects are integrated with each other, which are distance factor and fingerprint impression factor. The key aspect of propound mechanism is that the counterfeit online transaction cannot be performed without knowledge of the user. Graphical presentation of this mechanism is showing in the flowchart (as shown in Fig. 4) and portray in sequential steps as given below (Fig. 3):

Steps:

- Start legitimate Online Transaction.
- Financial application is opened by the user in cell phone with a predefined code.
- Through financial application, send the position of user's cell phone to the financial institution.
- Input details of transaction payment card to the online transaction instrument (i.e., PC or ATM or POS).
- Authenticate the details of transaction payment card and the gap between cell phone of user and transaction instrument by financial institution.

Fig. 3 Procedure of integrated approach of authentication and authorization propound mechanism

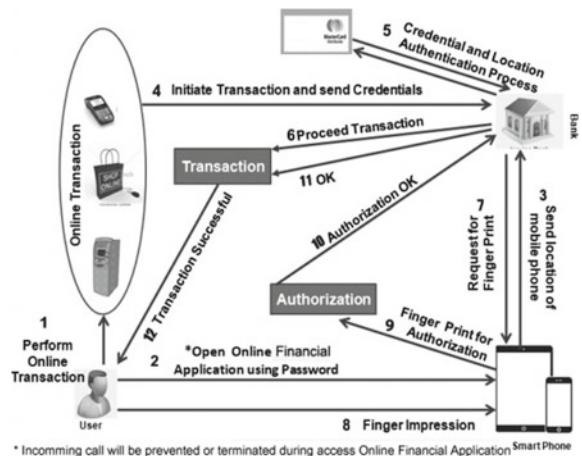
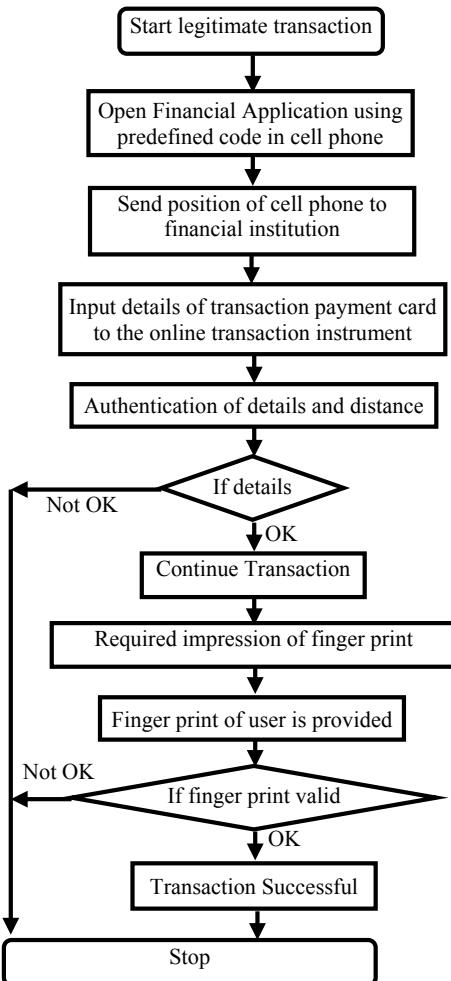


Fig. 4 Flowchart of integrated approach of authentication and authorization propound mechanism



6 If details OK.

```

then
  Continue Transaction
  Jump to sequence 7.
else
  Abort Transaction
  Jump to sequence 11.

```

- 7 Required impression of fingerprint of user through Financial Application.
- 8 Fingerprint of user is provided.
- 9 Check the validity of fingerprint of user by authorized process.

- 10 If fingerprint valid.
 - then Completion of successful transaction.
 - else Completion of unsuccessful transaction.
- 11 Stop

4 Cases of Counterfeit Online Transaction

The incidences of counterfeit online transaction are reported in various newspapers and media in which few cases are discussed in the subsequent sections. The effect of propound mechanism on counterfeit online transaction is also analyzed.

4.1 Counterfeit Online Transaction Fraud Using Phone

This event was taken place in the metro city of India in the year 2016. A fraudster who was acted as a delegate of financial institution made a call to the victim and informed about blocking of payment card. To unblock payment card, fraudster asked the user to verify the details of the payment card. The victim gave all the details of payment card to the fraudster. Actually, fraudster was performing counterfeit online transaction simultaneously. All the details which were acquired from the victim were inputted in online transaction simultaneously. For deduction of Rs. 5000/-, an OTP was sent to the victim. This OTP was also told by the victim as requested by the fraudster. The final outcome should be the deduction of amount but due to insufficient money in the account, no money was deducted. But a penalty charge was applied and deducted from the account of the victim as exposed in Fig. 5.

18.06.16 10000532502970616610013717CANC710.00	710.00	892.83Cr
25.06.16 INTEREST CREDIT	74.00	966.83Cr
05.07.16 INSUF BAL POS DECLINE CHARGE-050716	17.00	949.83Cr
TRANSFER TO 09835303228		

Fig. 5 Point of sale decline penalty charge due to deficient amount in the account

4.2 Counterfeit Online Transaction by Cloning of Payment Card

This event occurred in 2015 in one of the states (Uttar Pradesh) of India. A fraudster gave a skimmer device to the staff member of the restaurant and made a mutual consent that when any customer wanted to pay the bill through payment card, then the payment card should be swapped on skimmer device also and PIN number should be recorded. The same activity was carried out by the staff members. The details of the payment card and PIN number were delivered to the fraudster. Using these details, the fraudster made a clone of the card and used for the counterfeit online transaction. The outcome of this event was the deduction of amount from the account.

4.3 Effect of Propound Mechanism on Counterfeit Online Transaction Cases

The above discussed cases are considered and analyzed by the authors. The authors observed that fraudster stole all the necessary details of payment card including OTP of the user. In abstract, the actual existence user during the online transaction was not required. If propound integrated authentication and authorization mechanism is applied to the above discussed cases, then the counterfeit online transaction will not be executed because the physical presence of the legitimate user is required. This physical presence relates to the gap between cell phone and transaction instrument and also relates to fingerprint impression of the user which will not be possible to complete the requirement by the fraudster. This means that after stealing all the details, the fraudster cannot perform the counterfeit online transaction.

5 Conclusion with Future Work

For creating an unbeaten authentication and authorization mechanism to validate the online transaction, the authors investigated different types of counterfeit online transaction fraud. After analysis, the authors developed an integrated approach by combining authentication using distance factor and authorization using fingerprint impression. The propound mechanism reduced the counterfeit online transaction and increased the security of the valid online transaction. The principle thought behind this propound mechanism is to ensure that the only endorsed user is permitted to carry out a legitimate online transaction. If any fraudster tries to perform counterfeit online transaction, then at the initial authentication, i.e., using distance factor, online transaction will be stopped. If initial authentication is clear, then it will be stopped during authorization of fingerprint. The possibility of performing the counterfeit online transaction is very less. The authors also figured out the impact of propound

mechanism on various cases of counterfeit online transaction and concluded that the propound mechanism is suitable and useful to stop the counterfeit online transaction. In future, this propound mechanism will be implemented to check its validity and accuracy, and the outcome will be measured and compared with the current system to ensure the strength of propound mechanism.

Acknowledgements We extend our gratitude to the Integral University for acknowledging our research work and providing us with Manuscript Communication Number-IU/R&D/2018-MCN000408. We are also indebted to Shri Ramswaroop Memorial University for providing us the financial support for this research work.

References

1. Bhatla, T.P., Prabhu, V., Dua, A.: Understanding credit card frauds. *Cards Bus. Rev.* **1**(6) (2003)
2. Khattri, V., Singh, D.K.: A Novel Distance Authentication Mechanism to Prevent the Online Transaction Fraud. *Advances in Fire and Process Safety*, pp. 157–169. Springer, Singapore (2018)
3. Akole, P., Mane, N., Shinde, K., Swati, A.K.: Secure transactio: an credit card fraud detection system using visual cryptography. *Int. Res. J. Eng. Technol. (IRJET)*. **3**(4), 1719–1724 (2016)
4. Ogbanufe, O., Kim, D.J.: Comparing fingerprint-based biometrics authentication versus traditional authentication methods for e-payment. *Decis. Support Syst.* **106**, 1–14 (2018)
5. Kulat, A., Kulkarni, R., Bhagwat, N., Desai, K., Kulkarni, M.P.: Prevention of Online transaction frauds using OTP generation based on dual layer security mechanism. *Int. Res. J. Eng. Technol. (IRJET)*. **3**(4), 1058–1060 (2016)
6. Panigrahi, S., Kundu, A., Sural, S., Majumdar, A.K.: Credit card fraud detection: a fusion approach using dempster-shafer theory and bayesian learning. *Inf. Fusion.* **10**(4), 354–363 (2009)
7. Halvaaee, N.S., Akbari, M.K.: A novel model for credit card fraud detection using Artificial Immune Systems. *Appl. Soft Comput.* **24**, 40–49 (2014)
8. Soltani, N., Akbari, M.K., Javan, M.S.: A new user-based model for credit card fraud detection based on artificial immune system. In: 16th IEEE CSI International Symposium on Artificial Intelligence and Signal Processing, pp. 029–033 (2012)
9. Assis, C.A., Pereira,A., Pereira, M.A., Carrano, E.G.: A genetic programming approach for fraud detection in electronic transactions. In: IEEE Symposium on Computational Intelligence in Cyber Security, pp. 1–8 (2014)
10. RamaKalyani, K., UmaDevi, D.: Fraud detection of credit card payment system by genetic algorithm. *Int. J. Sci. Eng. Res.* **3**(7), 1–6 (2012)
11. Prakash, A., Chandrasekar, C.: An Optimized multiple semi-hidden Markov model for credit card fraud detection. *Indian J. Sci. Technol.* **8**(2), 165–171 (2015)
12. Khan, M.Z., Pathan, J.D., Ahmed, A.H.E.: credit card fraud detection system using hidden Markov model and k-clustering. *Int. J. Adv. Res. Comput. Commun. Eng.* **3**(2), 5458–5461 (2014)
13. Bekirev, A.S., Klimov, V.V., Kuzin, M.V., Shchukin, B.A.: Payment card fraud detection using neural network committee and clustering. *Opt. Memory Neural Netw.* **24**(3), 193–200 (2015)
14. Van Vlasselaer, V., Bravo, C., Caelen, O., Eliassi-Rad, T., Akoglu, L., Snoeck, M., Baesens, B.: APATE: a novel approach for automated credit card transaction fraud detection using network-based extensions. *Decis. Support Syst.* **75**, 38–48 (2015)
15. Khan, A.U.S., Akhtar, N., Qureshi, M.N.: Real-Time Credit-Card Fraud Detection using Artificial Neural Network Tuned by Simulated Annealing Algorithm. In: Proceedings of International Conference on Recent Trends in Information, Telecommunication and Computing, pp. 113–121 (2014)

16. Behera, T.K., Panigrahi, S.: Credit card fraud detection: a hybrid approach using fuzzy clustering & neural network. In: Second IEEE International Conference on Advances in Computing and Communication Engineering (ICACCE), pp. 494–499 (2015)
17. Credit Card and Debit Card Fraud Statistics, <https://wallethub.com/edu/credit-debit-card-fraud-statistics/25725/> (2018). Accessed on June 2018
18. Agoyi, M., Seral, D.: The use of SMS encrypted message to secure automatic teller machine. *Proc. Comput. Sci.* **3**, 1310–1314 (2011)
19. Sankhwar, S., Pandey, D.: A safeguard against ATM fraud. In: 6th IEEE International Conference on Advanced Computing (IACC), pp. 701–705 (2016)
20. Nti, I.K., Ansere, J.A., Appiah, A.: Investigating ATM Frauds In Sunyani Municipality: Customer's Perspective. *Int. J. Sci. Eng. Appl.* **6**(02), 59–65 (2017)
21. Sakharova, I.: Payment card fraud: Challenges and solutions. In: IEEE International Conference on Intelligence and Security Informatics, pp. 227–234 (2012)
22. [www.eMarketer.com:](http://www.eMarketer.com/) <https://apsalar.com/2016/03/mobile-and-smartphone-usage-statistics-for-india/> (2018). Accessed on June 2018

A Comparative Analysis of Transforms for Infrared and Visible Image Fusion



Apoorav Maulik Sharma, Renu Vig, Ayush Dogra, Bhawna Goyal and Sunil Agrawal

Abstract Image fusion is the art of combining two different images which are either captured on different times, using different sensors, from different focal points or from different modalities to fuse the best available within two into single one. The fusion of infrared and visible images has a widespread application in the field of military surveillance and night vision imaging technologies. The era of evolution of various transforms has led to the documentation of various efficient representational algorithms in literature, for instance, Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT) for the fusion of images. It is clearly stated in the field of image fusion that high quality of source images largely affects the image fusion rate. Therefore, in this paper, we explore and compare various transform-based image fusion techniques for noisy visible and infrared images.

Keywords Infrared · Visible · Multi-scale decomposition · DCT · Wavelet transform

1 Introduction

Image fusion has gained a lot of popularity due to its ability to represent the information in a more perceptible way to human beings. In image fusion, the complementary available information from two source images is fused together [1–4]. A general image fusion process has different stages. The first stage comprises InfraRed (IR) and visible images which are acquired using different sensors. The visible image contains the data which is acquired in the visible band of electromagnetic spectrum. It is normally perceptible to human eye. On the other hand, IR image contains the data which is acquired in the IR band of electromagnetic spectrum. It reveals the data which is normally not visible to human eye, or we can say it reveals the concealed features of the subject in the image. The IR sensors capture the IR radiations coming

A. M. Sharma (✉) · R. Vig · A. Dogra · B. Goyal · S. Agrawal
UIET, Panjab University, Chandigarh, India
e-mail: apoorav1207@gamil.cokm

from subject due to the heat present in the body of the subject. The IR images are captured in three sub-bands of IR band, namely, near infrared, mid-infrared, and far infrared. Images are captured in these three sub-bands according to the environment in which the images are being captured [5]. Hence, infrared and visible image fusion comes under the category of multi-sensor image fusion which is one of the categories of image fusion based on types of images. Other categories include multi-modal, multi-temporal, and multi-focus image fusion. Some major areas of IR and visible image fusion applications include civil and military surveillance [6], concealed weapon detection [7], remote sensing [8], and medical diagnosis [9–13]. In the literature, multiple data sets are used to validate the results of the algorithm.

In a typical image fusion process, transforms are applied to images in the next stage. The transform decomposes the images into different sub-bands which represent the high-pass and low-pass information present in the source images. This process is termed as multi-scale decomposition. At each scale of decomposition, resolution of the image is halved. Further, these sub-bands in the form of coefficients are fused together using some fusion rule in the next stage, and then, inverse transform is applied to retrieve back the fused image which is the final stage of a typical image fusion process [2].

1.1 Categories of Image Fusion

Generally, image fusion can be divided into three broad categories, viz.:

(a) Pixel-Level Image Fusion

When fusion is performed on the lowest level of an image, i.e., a pixel, then it is called pixel-level image fusion. Since it is performed on the most basic level of an image, this type of image fusion produces most accurate results.

(b) Feature-Level Image Fusion

In feature-level image fusion, fusion is done on the level of features. Hence, segmentation and feature extraction are important parts of it. The segmentation of input images is done using segmentation maps, and then, some activity-level measurements are done for each and every segmented region.

(c) Decision-Level Image Fusion

In this, fusion is done on the level of decision or symbol. So, when abstraction level is high, the fusion quality gets deteriorated [4].

2 Evolution of Transforms

Since long, numerous signal representational tools have been designed in order to construct the harmonic analysis of the input signal. The most former of these methods included the well-known Fourier transform. This enabled the mapping of the time domain information into the frequency domain analysis. The FT has been widely extended to represent two-dimensional signals; however, it has efficiently represented the smooth data but failed to represent the data with discontinuities and singularities. In other words, Fourier transform is only efficient in representing signals which do not change much over time and also, while representing a signal in frequency domain, some of its time domain information is lost. To address all these drawbacks, Short-Time Fourier Transform (STFT) was introduced in which a window was used to analyze a small portion of the signal at a time. It mapped the signal as a function of time and frequency simultaneously. The main drawback associated with STFT was its fixed window size. For a variety of signals available for analysis, a more flexible window size was required [13]. Later on, in the field of image processing, extensions of the Fourier transform were used to analyze the signal. One such extension was Discrete Cosine Transform (DCT). In image processing, DCT was used for lossy compression of the images as done by Joint Photography Expert Group (JPEG). In DCT, each pixel value of the image is transformed and represented as a sum of cosines in which we get the DC and AC coefficients. DCT for 2D signal can be represented mathematically as

$$(j_1, j_2) = \sqrt{\frac{4}{N^2}} \alpha(j_1)\alpha(j_2) \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} f(k_1, k_2) \cos\left(\frac{\pi(2k_1+1)j_1}{2N}\right) \cos\left(\frac{\pi(2k_2+1)j_2}{2N}\right), \quad (1)$$

where

$$\alpha(j) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } j = 0 \\ 1 & \text{for } j \neq 0 \end{cases} \quad \text{and } j_1, j_2, k_1, k_2 = 0, 1, 2, \dots, K-1.$$

The above equation is applicable when $k_1 = k_2$, i.e., when $f(k_1, k_2)$ is a square image with order $K \times K$ and when there is a rectangular image, then the limits of summation simply changes to the respective dimensions of the image. When $j_1 = j_2 = 0$, it represents the DC coefficient, and when $k_1 \neq k_2$, it represents the AC coefficients. Similarly, the expression of inverse DCT can be written as

$$f(k_1, k_2) = \sqrt{\frac{4}{N^2}} \alpha(j_1)\alpha(j_2) \sum_{k_1=0}^{K-1} \sum_{k_2=0}^{K-1} F(j_1, j_2) \cos\left(\frac{\pi(2k_1+1)j_1}{2N}\right) \cos\left(\frac{\pi(2k_2+1)j_2}{2N}\right). \quad (2)$$

To compute the DCT, image is considered in blocks of fixed size; then the above expression is first applied to that block. Generally, a 8×8 block is considered for the application of DCT. After transforming the image block, quantization operation is applied in which coefficients are quantized to the nearest integer values. The

coefficients corresponding to higher spatial frequencies become zero, and they can be discarded in order to compress the image. But, in image fusion, coefficients obtained are not quantized; rather, they are fused directly with the help of some appropriate fusion rule and then transformed back using IDCT. DCT has been used in image fusion effectively [14–17] but the problem associated with DCT is with higher spatial frequencies. Higher frequencies are associated with edges and boundaries present in the image. DCT blurs the image and fails to effectively represent the edges and contours.

The all above transforms posed the problem of local representation of the signals or signals with discontinuities. That is, where wavelets came into the picture. The wavelet analysis uses a waveform which is limited in time and has an average value of zero to transform the signals. The transform coefficients are obtained by doing the inner product of the signal and shifted and scaled versions of the mother wavelet. Mathematically, the wavelet transform of a 2D signal [13]:

$$W_{s,p} = \langle f, \psi_{s,p} \rangle, \quad (3)$$

where $\psi_{s,p}$ is the mother wavelet defined as

$$\psi_{s,p}(t) = 2^{-s} \psi\left(\frac{t - 2^s p}{2^s}\right). \quad (4)$$

And $W_{s,p}$ are the wavelet coefficients obtained on scale (s) and position (p) as

$$W_{s,p} = \int_{-\infty}^{\infty} f(t) \psi_{s,p}(t) dt. \quad (5)$$

And for reconstruction,

$$f(t) = \sum_{s,p} W_{s,p} \psi_{s,p}(t). \quad (6)$$

Here, both $s, p \in Z$. An array of high-pass and low-pass filters is used to decompose the signal. After the multi-scale decomposition of the signal, the approximate and detail sub-bands are obtained. There are different families of wavelet which can be used to decompose the signal for, e.g., haar, db (Daubechies), coiflets, symlets, bi-orthogonal wavelets, etc. Wavelets were efficient in representing the point singularities but they suffered from the problem of non-directionality. So, more directional version of wavelets came as Complex Wavelets (CWT) which was directional but they needed the development of some complex wavelets for decomposition which were not easy to design. Also, due to the complexity involved with the design of complex wavelets, it was computationally complex. So later on, Dual-Tree Complex Wavelet Transform (DTCWT) came as a solution to this problem [18].

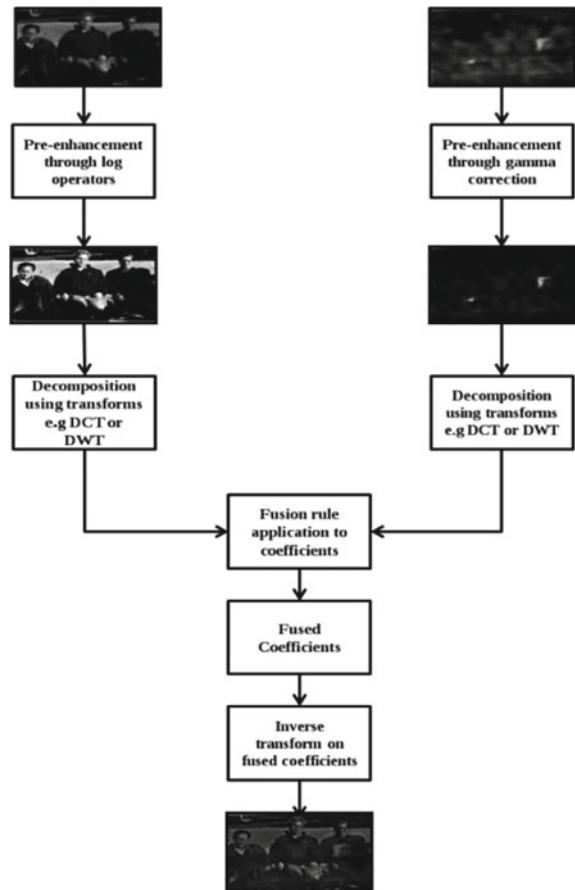
3 Proposed Methodology

The source images used in experimentation have low contrast so in order to enhance the details and contrast log and gamma operators are used; then, IR and the visible images are first decomposed separately as shown in Fig. 1, using a transform to obtain the transform coefficients. These transform coefficients are then fused using fusion rules. The two most widely used fusion rules are:

- (a) **Average fusion rule:** in this rule, the corresponding pixels of the two decomposed images are averaged together. Mathematically, it can be represented as

$$F(i, j) = 0.5 \times [I(i, j) + V(i, j)]. \quad (7)$$

Fig. 1 Flow diagram of proposed methodology



- (b) **Choose max fusion rule:** In this rule, maximum of the two values is chosen to represent the fused image. Mathematically, it can be represented as

$$F(i, j) = \max[I(i, j), V(i, j)]. \quad (8)$$

- (c) **Choose max fusion rule:** In this rule maximum of the two values is chosen to represent the fused image. Mathematically it can be represented as

$$F(i, j) = \max[I(i, j), V(i, j)], \quad (8)$$

where i, j are the pixel positions. I , V , and F are the infrared, visible, and fused image, respectively. After applying the appropriate fusion rule, fused image coefficients are obtained. These coefficients are then transformed back to the image using the inverse transform. A flow diagram of the proposed methodology is shown in Fig. 1.

There are certain parameters which are used to assess the quality of the fused image with respect to the source images. These parameters can be categorized as follows:

- (a) **Reference-based metrics:** These are the old tools used to evaluate the quality of the fused image for, e.g., entropy, standard deviation, spatial frequency, average gradient, mutual information, Structural Similarity Index Measure (SSIM), Root Mean Square Error (RMSE), correlation coefficient, Peak Signal-to-Noise Ratio (PSNR), etc. [1–3].
- (b) **Non-reference-based metrics:** These are the new metrics for in-depth analysis. Currently, these metrics are most widely used to assess the fused image quality:

$$Q^{STF} = \left(\frac{\sum_{i=1}^M \sum_{j=1}^N (Q^{SF}(i, j) \times \omega_S(i, j) + Q^{TF}(i, j) \times \omega_T(i, j))}{\sum_{i=1}^M \sum_{j=1}^N (\omega_S(i, j) + \omega_T(i, j))} \right), \quad (9)$$

where $Q^{ST}(i, j) = Q_\beta^{ST}(i, j)Q_\alpha^{ST}(i, j)$ and $Q_\beta^{ST}(i, j)$, $Q_\alpha^{ST}(i, j)$ are edge strength and orientation preservation values, and $\omega_A(i, j)$, $\omega_B(i, j)$ are weights to measure the importance of edge strength and orientation preservation values. The metric represented by Eq. 9 is the most widely used non-reference-based fusion evaluation metric.

4 Experiments and Results

Experiments are performed on the images named “gunIR.jpg” and “gunVIS.jpg”, representing the IR and Visible images, respectively (Fig. 2).

The above images are then fused together to get the fusion results. It can be noted that these images have low contrast and they contain noise. This adds to poor

visibility and hence poor visual perception. Hence, we used log operators to pre-enhance the images so that we can obtain better fusion results. The below images show the visual results of the fusion. Perception-wise there seems no much difference in the fused images; however, the evaluation metrics reveal the true picture of the fusion performance (Fig. 3).

All the DCT algorithms are implemented with block size 4 and 8. The DCTah technique used average fusion rule to fuse DC and the lowest AC components while the rest of the components were fused with choose max fusion rule. The DCTav technique used average fusion rule to fuse all the coefficients. In DCTch, the lower AC and the DC components are fused using average fusion rule but the higher AC components were chosen according to contrast (Table 1).

Finally, in DCTe, average fusion rule and energy criteria were used to choose the coefficients for the fused image. It can be seen that Discrete Cosine Harmonic Wavelet Transform (DCHWT) provide the best fusion results overall in comparison to other transforms (as shown by highlighted parameters). Also, it is very interesting to note here that the DCTe technique is not only close in performance but it has low value of information loss. However, the key point to note here is that the pre-enhancement procedures performed before fusion serve the purpose of improving the human perceptibility of fused images as the results are perceptually better. Also, objectively on evaluation parameters, results are better than the techniques discussed in [16, 17].

Fig. 2 Source images

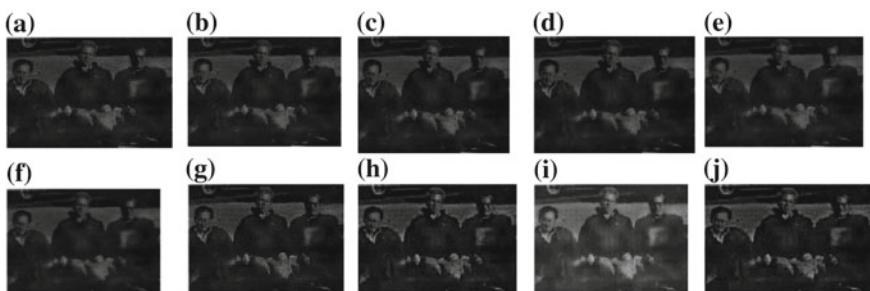
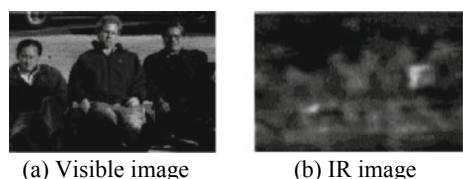


Fig. 3 Image fusion using **a** DCTah4, **b** DCTah8, **c** DCTav4, **d** DCTav8, **e** DCTch4, **f** DCTch8, **g** DCTe4, **h** DCTe8, **i** DWT, and **j** DCHWT

Table 1 Objective evaluation metrics for fusion performance

Technique	Q^{STF}	L^{STF}	Metrics			
DCTah4	0.8293	0.1564	SD	AG	E	MI
DCTah8	0.8168	0.1692	30.4415	33.1239	6.0205	1.8367
DCTav4	0.6769	0.3231	25.9028	18.2055	6.0507	3.8840
DCTav8	0.6769	0.3231	25.9028	18.2055	6.0528	3.8481
DCTch4	0.6757	0.3086	28.8743	27.8562	6.0888	1.7175
DCTch8	0.6978	0.2871	29.1795	29.5494	6.0680	1.7329
DCTe4	0.8642	0.1072	33.3054	34.3906	5.9197	1.9300
DCTe8	0.8675	0.0964	36.1742	34.1912	5.8676	1.8619
DWT	0.6559	0.3407	27.7770	24.6807	6.1335	1.6202
DCHWT	0.8804	0.1151	37.2681	29.5047	6.5207	2.2075

5 Conclusions

It can be concluded from the above discussion that although these transform or decomposition methods are primitive, they still hold a strong position in this field of research. In literature, they are still used in a hybrid fashion to improve the quality of the fusion process. Due to some drawbacks, other transforms like curvelets and contourlets were derived out but these transform still hold a respectable position in the area of image processing. However, it is important to note here that despite having considerable performance in terms of evaluation metrics, a lot of work is needed to be done to improve the human perception quality of the image. For the fusion of images with inherent noise, some pre-enhancement procedures can be followed to improve the visual quality of the fused image.

References

1. Ma, J., Ma, Y., Li, C.: Infrared and visible image fusion methods and applications: a survey. *Inf. Fusion* **45**, 153–178 (2019)
2. Li, S., Kang, X., Fang, L., Hu, J., Yin, H.: Pixel-level image fusion: A survey of the state of the art. *Inf. Fusion* **33**, 100–112 (2017)
3. James, A.P., Dasarathy, B.V.: Medical image fusion: A survey of the state of the art. *Inf. Fusion* **19**, 4–19 (2014)
4. Dogra, A., Goyal, B., Agrawal, S.: From multi-scale decomposition to non-multi-scale decomposition methods: a comprehensive survey of image fusion techniques and its applications. *IEEE Access* **5**, 16040–16067 (2017)
5. Waxman, A.M., Gove, A.N., Fay, D.A., Racamato, J.P., Carrick, J.E., Seibert, M.C., Savoye, E.D.: Color night vision: opponent processing in the fusion of visible and IR imagery. *Neural Netw.* **10**(1), 1–6 (1997)
6. Toet, A.: Iterative guided image fusion. *Peer J.Comput. Sci.* **2**, e80 (2016)

7. Kumar, B.S.: Image fusion based on pixel significance using cross bilateral filter. *SIViP* **9**(5), 1193–1204 (2014)
8. Ghasseman, H.: A review of remote sensing image fusion methods, *Inf. Fusion* **32**, 75–89 (2016)
9. Dogra, A., Goyal, B., Agrawal, S., Ahuja, C.: K: Efficient fusion of osseous and vascular details in wavelet domain. *Pattern Recogn. Lett.* **94**, 189–193 (2017)
10. Dogra, A., Agrawal, S., Goyal, B., Khandelwal, N., Ahuja, C.K.: Color and grey scale fusion of osseous and vascular information. *J. Comput. Sci.* **17**, 103–114 (2016)
11. Dogra, A., Goyal, B., Agrawal, S.: Current and future orientation of anatomical and functional imaging modality fusion. *Biomed. Pharmacol. J.* **10**(4), 1661–1663 (2017)
12. Zheng, Y.: Image Fusion and its applications (2011)
13. Misiti, M., Misiti, Y., Oppenheim, G., Michel, J.P.: Wavelet toolbox: for use with MATLAB (1996)
14. Naidu, V.P.S.: Discrete cosine transform-based image fusion. *Def. Sci. J.* **60**(1), 48–54 (2010)
15. Kumar, B.S.: Multifocus and multispectral image fusion based on pixel significance using discrete cosine harmonic wavelet transform. *SIViP* **7**(6), 1125–1143 (2013)
16. Paramanandham, N., Rajendiran, K.: Infrared and visible image fusion using discrete cosine transform and swarm intelligence for surveillance applications. *Infrared Phys. Technol.* **88**, 13–22 (2018)
17. Jin, X., Jiang, Q., Yao, S., Zhou, D., Nie, R., Lee, S.J., He, K.: Infrared and visual image fusion method based on discrete cosine transform and local spatial frequency in discrete stationary wavelet transform domain. *Infrared Phys. Technol.* **88**, 1–12 (2018)
18. Kingsbury, N.: Rotation-invariant local feature matching with complex wavelets. In: 2006 14th European Signal Processing Conference, pp 1–5. IEEE (2006)

Selection of Optimal Performance Parameters of Alumina/Water Nanofluid Flow in Ribbed Square Duct by Using AHP-TOPSIS Techniques



Sunil Kumar, Anil Kumar, Alok Darshan Kothiyal and Mangal Singh Bisht

Abstract In the present article AHP, entropy, and TOPSIS techniques are applied to select the optimal performance parameters of alumina/water nanofluid flow in protrusion ribbed square channel. AHP and entropy method is used to determine the subjective, objective, and synthesis weights of all three PDAs which are Nu_{ave} , f_{ave} , and η_p respectively. TOPSIS technique is used to find relative closeness index and to rank all the alternatives which offer the optimal alternative. The alternatives with a maximum value of closeness index are nominated as an optimal alternative. From the results, alternative order based on relative closeness index is A15 > A16 > A14 > A7 > A11 > A8 > A6 > A12 > A9 > A13 > A10 > A5 > A4 > A3 > A2 > A1. The alternative A15 having maximum value 0.8459 of closeness index exhibits optimal performance.

Keywords AHP-TOPSIS · Optimal parameter · Nanofluid · Flow channel

1 Introduction

Energy is the critical constituent of the world which simplifies life development and sustainability to its active residents [1]. With the increase in alertness of energy scarcity, energy saving is essential for cost-effective, ecological, and societal reasons and to offer numerous benefits including reduced air pollution and fuel consumption [2]. The enhancement of the heat transfer proportions in the thermal means by

S. Kumar (✉)
Department of Mathematics, UTU, Dehradun, India
e-mail: reachtome.sunil@gmail.com

S. Kumar · A. Kumar
School of Mechanical & Civil Engineering, Shoolini University, Solan, India

A. D. Kothiyal
Department of Mathematics, Doon College Rishikesh, Rishikesh, Uttarakhand, India

M. S. Bisht
Department of Basic Sciences, GBPE College, New Delhi, Uttarakhand, India

implementing appropriate method can effect in noteworthy methodical compensations and price savings. [3, 4]. The heat exchangers augmented with different geometrical factors like Nu_{ave} , f_{ave} , and η_p are originate to be appropriate attributes to forecasting the enactment of ribbed duct [5, 6].

A modest approach of selecting optimal parameters for heat transfer efficiency of the system needs careful selection of PDAs [7]. Hence, MADM model is to be selected which provides the best alternative solution to optimize these PDAs. In engineering design, several methods are applied to choose the superlative substitute in the primary phase of design course such as AHP [8], FAHP (fuzzy analytical hierarchy process) [9], FSAW (simple additive weighting) [10], GRA (grey relational analysis) [11], PROMETHEE(II) (preference ranking organization method for enrichment evaluation) and TOPSIS (technique for preference by similarity to the ideal solution) [12, 13], RSM (response surface methodology) [14]. Taylor et al. [15] implemented the hybrid type of FAHP and FTOPSIS approach for selection of key peril criterion of their projects. Zouggari and Benyoucef [16] applied AHP-TOPSIS technique, while Sun [17] made use of these approaches to improve an assessment of model for their methodical decision-making tool.

The objective of this article is to implement AHP, entropy, and TOPSIS approaches for selecting optimal parameters for ribbed square duct.

2 Experimental Details and Range of Parameters

An experimental setup was designed and made-up according to ASHRAE Standard [18]. The full details of experiments had been presented in the previous accepted article, readers may follow; Kumar et al. [19] “Effect of nanofluid flow and protrusion transverse ribs on thermal and hydrodynamic performance in square channel: An experimental investigation” Journal of Enhanced Heat Transfer, 2018.

3 Methodology

The methodology followed in this research article contains three key phases discussed below.

Phase-1: Determination of Alternatives and Attributes

In the first phase the alternatives and PDAs are determined, that act as judgment variable for evaluation of performance of square ribbed duct. The same are given in Table 1.

Phase-2: AHP and Entropy Method for Weight Evaluation

AHP Method

In AHP first we define the problem goal, and then the goal is decomposed into comparative judgment of the attributes and the alternatives. The set of n attributes compared pairwise in accordance with importance weights.

Table 1 Experimental data for different operating parameters and performance defining attributes

Alternatives	Operating parameters				Attributes		
	Re_n	X_s/d_p	Y_s/d_p	e_p/d_p	Nu_{ave}	f_{ave}	η_p
A1	6000	1.8	1.8	1.67	72	0.161	1.8845
A2	8000	1.8	1.8	1.67	94.4	0.148	1.9910
A3	10000	1.8	1.8	1.67	116.5	0.137	2.0726
A4	12000	1.8	1.8	1.67	136	0.127	2.1452
A5	14000	1.8	1.8	1.67	155.5	0.118	2.1982
A6	16000	1.8	1.8	1.67	174	0.109	2.2480
A7	18000	1.8	1.8	1.67	187	0.101	2.2609
A8	18000	1.4	1.8	1.67	180	0.098	2.2261
A9	18000	2.2	1.8	1.67	172.5	0.091	2.1998
A10	18000	2.6	1.8	1.67	163.8	0.087	2.1803
A11	18000	1.8	1.4	1.67	180	0.091	2.2361
A12	18000	1.8	2.2	1.67	172.5	0.085	2.2128
A13	18000	1.8	2.6	1.67	163.8	0.079	2.1903
A14	18000	1.8	1.8	0.83	196	0.109	2.2899
A15	18000	1.8	1.8	1	217	0.132	2.3749
A16	18000	1.8	1.8	1.25	202	0.118	2.3199

The rating scale employed to relate the weight importance among attributes according to the linguistic meaning is from 1 to 9. The digit 1 is for similar importance, 3 for moderate importance, 5 for intense importance, 7 for demonstrated importance, and 9 for extreme importance. Whereas 2, 4, 6, and 8 digits are used to simplify somewhat differing verdicts or intermediate. The relative weights are found by calculating eigenvector (w) with respect to λ_{\max} which satisfy the equation

$$Aw = \lambda_{\max}w \quad (1)$$

Moreover, to confirm the consistency and accuracy of the relative weights, two indices, the consistency index ($C.I.$) is calculated using the equation

$$C.I. = (\lambda_{\max} - 1)/(n - 1) \quad (2)$$

Whereas C.R. is calculated as: $C.R. = C.I./R.I.$, here R.I. refers to a random consistency index. As in the present investigation the comparison matrix size is 3, so the value of R.I. given by Saaty [20] is 0.58.

Entropy Method

To calculate the weights by this method, the normalized decision matrix P_{ij} is obtained as

$$P_{ij} = X_{ij} / \sqrt{\sum_{i=1}^m X_{ij}^2} \quad (3)$$

Then entropy value E_j for j th attributes is calculated as:

$$E_j = -k \sum_{i=1}^m P_{ij} \ln P_{ij} \quad (4)$$

where $j = 1, 2, \dots, n$ and $k = 1 / \ln m$.

The degree (d_j) of each attribute is obtained as

$$d_j = |1 - E_j| \quad (5)$$

Then entropy weight of each attribute is calculated as

$$\beta_j = d_j / \sum_{j=1}^n d_j \quad (6)$$

Phase-3: TOPSIS Method

The different steps involved in this method are as follows:

Step I. In the first step the decision matrix is normalized using the equation

$$N_{ij} = X_{ij} / \sqrt{\sum_{i=1}^m X_{ij}^2}, i = 1, 2, \dots, m, j = 1, 2, \dots, n \quad (7)$$

Step II. In the second step the weighted normalized matrix is obtained as

$$V_{ij} = W_j \times N_{ij} \quad (8)$$

Step III. The positive ideal solutions (V^+) and negative ideal solutions (V^-) are calculated as

$$V_1^+, V_2^+, \dots, V_n^+ = \begin{cases} \max_i V_{ij}, & \text{if } j \text{ is benefit attributes} \\ \min_i V_{ij}, & \text{if } j \text{ is cost attributes} \end{cases} \quad (9)$$

$$V_1^-, V_2^-, \dots, V_n^- = \begin{cases} \max_i V_{ij}, & \text{if } j \text{ is benefit attributes} \\ \min_i V_{ij}, & \text{if } j \text{ is cost attributes} \end{cases} \quad (10)$$

Step IV. In this step Euclidean distance is calculated as

$$S_i^+ = \sqrt{\sum_{i=1}^m (V_{ij} - V_i^+)^2} \quad i = 1, 2, \dots, m, j = 1, 2, \dots, n \quad (11)$$

$$S_i^- = \sqrt{\sum_{i=1}^m (V_{ij} - V_i^-)^2} \quad i = 1, 2, \dots, m, j = 1, 2, \dots, n \quad (12)$$

Step V. Finally the relative closeness index (C_i) of the ideal solution is calculated as

$$C_i = \frac{S_i^-}{S_i^- + S_i^+}; \quad i = 1, 2, \dots, m; \quad 0 \leq C_i \leq 1 \quad (13)$$

4 Results and Discussion

4.1 Effect of Operating parameters on PDAs

Experimental results of operating parameters on PDAs are given in Table 1.

The values of all three performance defining attributes are presented graphically in Fig. 1. Figure 1a represents the effect of operating parameters at PDA-1 (Nu_{ave}) which indicated that value of PDA-1 is minimum (least performance) for A1 when operating parameters $Re_n = 6000$, $X_s/d_p = 1.8$, $Y_s/d_p = 1.8$ and $e_p/d_p = 1.67$ and maximum (optimal performance) for A15 when operating parameters $Re_n = 18000$, $X_s/d_p = 1.8$, $Y_s/d_p = 1.8$, and $e_p/d_p = 1.0$. Figure 1b displays the effect of operating parameters at PDA-2 (f_{ave}) which indicated that the value of PDA-2 is minimum for A13 when operating parameters $Re_n = 18000$, $X_s/d_p = 1.8$, $Y_s/d_p = 2.6$ and $e_p/d_p = 1.67$ whereas maximum for A1 when operating parameters $Re_n = 6000$, $X_s/d_p = 1.8$, $Y_s/d_p = 1.8$ and $e_p/d_p = 1.67$. Similarly Fig. 1c illustrates the effect of operating parameters at PDA-3 and showed that value of PDA-3 is minimum for A1 when operating parameters $Re_n = 6000$, $X_s/d_p = 1.8$, $Y_s/d_p = 1.8$, and $e_p/d_p = 1.67$ and maximum for A15 when operating parameters $Re_n = 18000$, $X_s/d_p = 1.8$, $Y_s/d_p = 1.8$, and $e_p/d_p = 1.0$. The above results noticeably indicated that different PDAs have dependence on the operating parameter configurations, which make the performance prediction trends more complicated. So to overcome this situation, the hybrid approach of AHP, entropy, and TOPSIS has been applied for optimization.

4.2 Determination of Weight for Different PDAs by AHP and Entropy Method

The subjective weight (α_j) is calculated using pairwise comparisons of PDAs given in Table 2. From pairwise comparison it is found that PDA-3 is the most important attribute. The value of (λ_{max}) calculated from pairwise comparison process was 3.0388. To determine the reliability of the comparison matrix, the calculated value

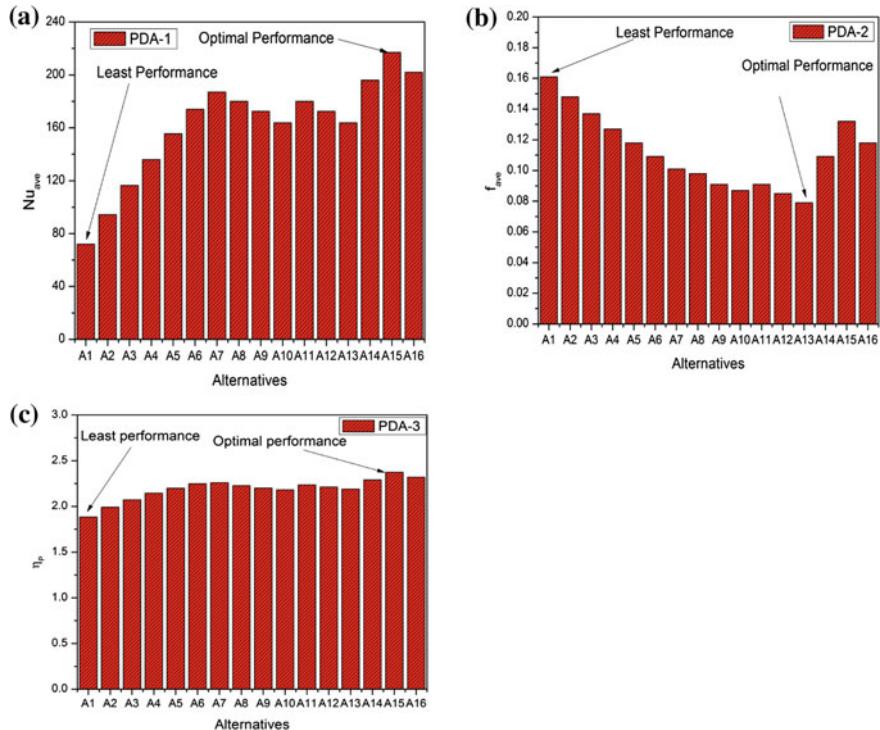


Fig. 1 **a** Variation of PDA-1 with different alternatives. **b** Variation of PDA-2 with different alternatives. **c** Variation of PDA-3 with different alternatives

Table 2 The pairwise comparison matrix for different PDAs

	PDA-1	PDA-2	PDA-3
PDA-1	1	3	0.333
PDA-2	0.333	1	0.2
PDA-3	3	5	1

of C.I. and C.R. are $0.0194 < 1$ and $0.0334 < 1$, respectively, which shows that comparison matrix is reliable. After calculating (α_j), the objective weight (β_j) is calculated. For calculating the value of (β_j) firstly the entropy value (E_j) and its degree (d_j) is calculated by using Eqs. (4) and (5). The values of $E_j = (1.9423, 1.9586, 1.9977)$ and $d_j = (0.9423, 0.9586, 0.9977)$ are for each PDAs respectively. The value of β_j calculated using Eq. (6) for each PDAs are shown in Table 4. From the values of (α_j) and β_j the value of synthesis weight (W_j) for each attributes given in Table 4 is calculated using Eq. (14).

$$W_j = (\alpha_j \times \beta_j) / \sum_{j=1}^n \alpha_j \times \beta_j \quad (14)$$

Table 3 Normalized and weighted normalized matrix

Alternatives	Normalized matrix			Weighted normalized matrix		
	PDA-1	PDA-2	PDA-3	PDA-1	PDA-2	PDA-3
A 1	0.1086	0.3521	0.2149	0.0272	0.0366	0.1386
A 2	0.1423	0.3237	0.2270	0.0357	0.0337	0.1464
A 3	0.1757	0.2997	0.2363	0.0441	0.0312	0.1524
A 4	0.2051	0.2778	0.2446	0.0515	0.0289	0.1578
A 5	0.2345	0.2581	0.2507	0.0588	0.0268	0.1617
A 6	0.2624	0.2384	0.2563	0.0659	0.0248	0.1653
A 7	0.2820	0.2209	0.2578	0.0708	0.0230	0.1663
A 8	0.2714	0.2133	0.2538	0.0681	0.0222	0.1637
A 9	0.2601	0.1990	0.2508	0.0653	0.0207	0.1618
A 10	0.2470	0.1903	0.2486	0.0620	0.0198	0.1604
A 11	0.2714	0.1990	0.2550	0.0681	0.0207	0.1645
A 12	0.2601	0.1859	0.2523	0.0653	0.0193	0.1627
A 13	0.2470	0.1728	0.2498	0.0620	0.0180	0.1611
A 14	0.2955	0.2384	0.2611	0.0742	0.0248	0.1684
A 15	0.3272	0.2887	0.2708	0.0821	0.0300	0.1747
A 16	0.3046	0.2581	0.2645	0.0764	0.0268	0.1706

4.3 Determination of Ranking of the Alternatives by TOPSIS Method

In TOPSIS firstly the decision matrix is normalized using Eq. (7) and then using Eq. (8) weighted normalized matrix is formed given in Table 3. The value of positive ideal solutions (V^+) and negative ideal solutions (V^-) of each PDAs determined from weighted normalized matrix using Eqs. (9), (10) are given in Table 4. After finding V^+ and V^- the Euclidean distances S_i^+ and S_i^- calculated for all alternatives using Eqs. (11), (12) are given in Table 5. Lastly, relative closeness index C_i is calculated to the ideal solution of all alternatives using Eq. (13) and results are presented in Table 5. In the end after the whole calculation the ranking of alternatives done to get optimal set of operating parameters is presented in Table 5. The alternatives with maximum value of C_i is nominated as an optimal alternative. Table 5 clearly shows that alternative A15 having value of operating parameters as $Re_n = 18000$, $X_s/d_p = 1.8$, $Y_s/d_p = 1.8$, and $e_p/d_p = 1.0$ have maximum value 0.8459 of C_i so ranked as 1 followed by alternative A16 having 0.8410 value of C_i . Alternative A1 has minimum value of C_i and is ranked at position 16. According to the value of C_i the ranking of alternatives in decreasing order is A15 > A16 > A14 > A7 > A11 > A8 > A6 > A12 > A9 > A13 > A10 > A5 > A4 > A3 > A2 > A1.

Table 4 Attribute weights and positive (V^+) negative (V^-) ideal solutions

	α_j	β_j	w_j	V^+	V^-
PDA-1	0.261	0.325	0.251	0.0821	0.0272
PDA-2	0.106	0.331	0.104	0.0180	0.0366
PDA-3	0.633	0.344	0.645	0.1747	0.1386

Table 5 Euclidean distances, closeness coefficient and ranking of the alternatives

Alternatives	S_i^+	S_i^-	C_i	Rank
A 1	0.0683	0.0000	0.0008	16
A 2	0.0565	0.0119	0.0119	15
A 3	0.0460	0.0225	0.3286	14
A 4	0.0367	0.0319	0.4651	13
A 5	0.0281	0.0404	0.5897	12
A 6	0.0199	0.0485	0.7084	7
A 7	0.0150	0.0534	0.7811	4
A 8	0.0183	0.0501	0.7330	6
A 9	0.0214	0.0214	0.6889	9
A 10	0.0248	0.0443	0.6416	11
A 11	0.0175	0.0509	0.7439	5
A 12	0.0207	0.0483	0.7001	8
A 13	0.0243	0.0454	0.6516	10
A 14	0.0122	0.0569	0.8236	3
A 15	0.0120	0.0660	0.8459	1
A 16	0.0113	0.0595	0.8410	2

5 Conclusions

This work explored the applicability of AHP-TOPSIS method, to selecting optimal performance parameter of alumina/water nanofluid flow in protrusion ribbed square channel to maximize the thermal as well as the hydraulic performance of the system. The following major conclusions are drawn from the study:

1. The AHP-TOPSIS method was proved to be a rapid and vigorous technique for assessment of the optimal design of alumina/water nanofluid flow in protrusion ribbed square channel parameters.
2. The alternatives with maximum value of C_i is nominated as an optimal alternative. The alternative order based on relative closeness index is A15 > A16 > A14 > A7 > A11 > A8 > A6 > A12 > A9 > A13 > A10 > A5 > A4 > A3 > A2 > A1. The alternative A15 having value of operating parameters as $Re_n = 18000$, $X_s/d_p = 1.8$, $Y_s/d_p = 1.8$, and $e_p/d_p = 1.0$ has maximum value 0.8459 of C_i and exhibits optimal performance.

References

1. Panos, E., Densing, M., Volkart, K.: Access to electricity in the World Energy Council's global energy scenarios: an outlook for developing regions until 2030. *Energy Strat. Rev.* **9**, 28–49 (2016)
2. Guo-Yan, Z., En, W., Shan-Tung, T.: Techno-economic study on compact heat exchangers. *Int. J. Energy Res.* **32**, 1119–1127 (2008)
3. Manca, O., Nardini, S., Ricci, D.: A numerical study of nanofluid forced convection in ribbed channels. *Appl. Therm. Eng.* **37**, 280–292 (2012)
4. Lauriat, G.: On the uses of classical or improved heat transfer correlations for the predictions of convective thermal performances of water-Al₂O₃ nanofluids. *Appl. Therm. Eng.* **129**, 1039–1057 (2018)
5. Kumar, S., Kothiyal, A.D., Bisht, M.S., Kumar, A.: Turbulent heat transfer and nanofluid flow in a protruded ribbed square passage. *Results Phys.* **7**, 3603–3618 (2017)
6. Kumar, S., Kothiyal, A.D., Bisht, M.S., Kumar, A.: Numerical analysis of thermal hydraulic performance of Al₂O₃-H₂O nanofluid flowing through a protrusion obstacles square mini channel. *Case Stud. Therm. Eng.* **9**, 108–121 (2017)
7. Chamoli, S.: ANN and RSM approach for modeling and optimization of designing parameters for a V down perforated baffle roughened rectangular channel. *Alex. Eng. J.* **54**, 429–446 (2015)
8. Chu, J., Su, Y.: The application of TOPSIS method in selecting fixed seismic shelter for evacuation in cities. *Syst. Eng. Procedia* **3**, 391–397 (2012)
9. Ayag, Z., Özdemir, R.G.: A fuzzy AHP approach to evaluating machine tool alternatives. *J. Intell. Manuf.* **17**, 179–190 (2006)
10. Chou, S.Y., Chang, Y.H., Shen, C.Y.: A fuzzy simple additive weighting system under group decision-making for facility location selection with objective/subjective attributes. *Eur. J. Oper. Res.* **189**, 132–145 (2008)
11. Chamoli, S., Yu, P., Kumar, A.: Multi-response optimization of geometric and flow parameters in a heat exchanger tube with perforated disk inserts by Taguchi grey relational analysis. *Appl. Therm. Eng.* **103**, 1339–1350 (2016)
12. Çalışkan, H., Kurşuncu, B., Kurbanoglu, C., Güven, Ş.Y.: Material selection for the tool holder working under hard milling conditions using different multi criteria decision making methods. *Mater. Des.* **45**, 473–479 (2013)
13. Chauhan, R., Singh, T., Tiwari, A., Patnaik, A., Thakur, N.S.: Hybrid entropy – TOPSIS approach for energy performance prioritization in a rectangular channel employing impinging air jets. *Energy* **134**, 360–368 (2017)
14. Singh, S., Dhiman, P.: Thermal and thermohydraulic performance evaluation of a novel type double pass packed bed solar air heater under external recycle using an analytical and RSM (response surface methodology) combined approach. *Energy* **72**, 344–359 (2014)
15. Taylan, O., Bafail, A.O., Abdulaal, R.M.S., Kabli, M.R.: Construction projects selection and risk assessment by fuzzy AHP and fuzzy TOPSIS methodologies. *Appl. Soft Comput.* **17**, 105–116 (2014)
16. Zouggari, A., Benyoucef, L.: Simulation based fuzzy TOPSIS approach for group multi-criteria supplier selection problem. *Eng. Appl. Artif. Intell.* **25**, 507–519 (2012)
17. Sun, C.C.: A performance evaluation model by integrating fuzzy AHP and fuzzy TOPSIS methods. *Expert Syst. Appl.* **37**, 7745–7754 (2010)
18. American Society of Heating Refrigerating and Air-Conditioning Engineers Inc. ASHRAE Handbook, American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc. (2015)
19. Kumar, S., Kothiyal, A.D., Bisht, M.S., Kumar, A.: Effect of nanofluid flow and protrusion transverse ribs on thermal and hydrodynamic performance in square channel: an experimental investigation. *J. Enhanc. Heat Transf.* (2018). <https://doi.org/10.1615/jenhheattransf.2018026042>
20. Saaty, T.L.: The Analytic Hierarchy process. McGraw-Hill, New York (1980)

A Review: Soil Moisture Estimation Using Different Techniques



Jitender Pandey, Vivek Chamoli and Rishi Prakash

Abstract In this research paper, we have discussed about the different methods for estimation of soil moisture. Soil moisture is an imperative segment from a socioeconomical point of view. The importance and application of the soil moisture is seen in different fields like agriculture, forecasting of drought, geological application, etc. In this paper, we went over the points of interest and impediments of the different techniques to be examined and endeavor to dissect the best method for deciding the soil moisture content.

Keywords Soil moisture · Hydrological · Remote sensing · Time domain reflectometry · Frequency domain reflectometry · Global navigation satellite system (GNSS)

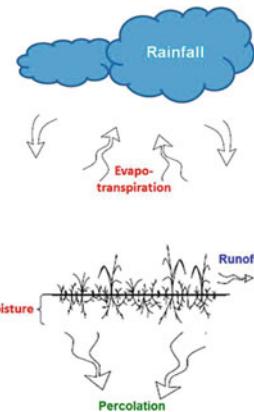
1 Introduction

Determining the moisture of the soil is one of the important things for various applications for example in the case of large agricultural fields. The moisture content can be used as raw data for various applications like determining the health of the soil. Many methods have been put to arrive at a simple and accurate way of determining soil moisture. There are many types of ways for estimating the moisture of soil. Techniques like time domain reflectometry method are relatively faster. Other methods like gravimetric techniques are somehow destructive, more time consuming, and laborious. However, the important thing is the spatial distribution of soil moisture which is used as an important parameter for various applications like metrological applications, agricultural applications, hydrological applications, forestry applications, monitoring of flood and drought. Thus we need an upscale point measurement

J. Pandey (✉) · V. Chamoli · R. Prakash
Graphic Era (Deemed to Be University), Dehradun, India
e-mail: jitenderpandey999@gmail.com

V. Chamoli
e-mail: vivekchamoli08@gmail.com

Fig. 1 Schematic of hydrological cycle [1]



to spatial scale such as to hydrological or fields. Some of the methods for determining the soil moisture are discussed in the upcoming sections.

1.1 Soil Moisture Estimation from Hydrological Models

The schematic representation of the hydrological cycle is shown in Fig. 1. The figure indicates the importance of the soil moisture in the entire hydrological cycle. The hydrological cycle's land phase is based on the water balance equation given in Eq. (1).

$$SWf = SWi + \sum_{i=1}^t (Rday - Qsurf - Ea - Wseep - Qgw) \quad (1)$$

Here SWi indicates initial soil moisture content, SWf is final soil water content, time is defined as t in days, i symbolizes amount of precipitation in a day, Qsurf the amount of surface runoff on day i, Ea stands for amount of evaporation, Seep stands for amount of percolation and bypass flow exiting the soil profile bottom, and Qgw stands for amount of return flow on day i. Figure 1 shows the schematic of hydrological cycle.

Estimation of soil moisture by hydrological methods deals with the computation of evapotranspiration, portioning between infiltration between runoff and infiltration, computation of lateral and vertical redistribution, and the numbers of soil layers which are used. In this approach for soil moisture estimation, two problems arose. First one is that assigning hydrological parameters to each subunit is somehow done in more or less arbitrary way. The second problem is that the interaction between the subunits can be assessed by the addition of those quantities which do not represent the ground reality.

Although the hydrologic method has the capability of providing timely information on the spatial moisture distribution without visiting the field. But the problem is that one needs to be cautious about the general disadvantage of error associated with their estimation which arose due to model error and observation error, i.e., structure error and parameter uncertainty error. These errors are, however, due to large spatial and temporal variation in the moisture of the soil. These all are arising due to the heterogeneity of the properties of soil, evaporation, precipitation, and vegetation [1, 2].

1.2 Dielectric Techniques

By this technique, the moisture content in the soil can be measured by utilizing the dielectric property of soil since there is a big difference between the pure water and dielectric constant of dry soil. Capacitance technique, TDR, frequency domain reflectometry are some of those major techniques which are used for measuring the soil moisture. There is a minimum influence of temperature on the measurement of electrical permittivity and because of this method measures the moisture content accurately. This type of technique can be further categorized in the following.

1.2.1 Time Domain Reflectometry Technique

In this technique, the TDR sensors measures the time which is needed for traveling of the transmitted signal from one end to the other end. The propagation constant for electromagnetic waves in the soil like attenuation and velocity depends on the property of the soil and used to determine the dielectric permittivity of the soil from the velocity of an EM wave which is emitted by the pulse generator and then passed along the waveguide of the probe of TDR. The dielectric constant is then used for the measurement of soil moisture content. The moisture content of the soil determined by this method is independent of the temperature, texture, and salinity of the soil. The disadvantage of this technique is that it is too expensive. Figure 2 shows the figure of TDR sensor [3, 4].

1.2.2 Frequency Domain Reflectometry Technique (FDR)

The principle of working of FDR is same as TDR probe. But the FDR tool measures the change in the frequency of the transmitted signal which can be altered by permittivity or by the moisture content of the soil in the ground. The probe is put inside the soil and then the electric field is applied which gives the reading because of the capacitance effect. Figure 3 shows the frequency domain reflectometry sensor.

When the GPS satellite rises above the horizon it sends out the signals. Most of the satellites rise and get set at different azimuths angles with respect to the antenna at the receiving section. The SNR data obtained from different satellite track is referred

Fig. 2 TDR sensor (<http://www.vanwalt.com/tdr-soil-moisture-measurement.html>)



Fig. 3 Frequency domain reflectometry sensor (<http://www.geo.uu.nl>)



by a number sequence called a pseudo-random number (PRN) and a quadrant of which the reflections from the ground came. When the satellite is at low elevation angle both the direct signal as well as indirect signal gets suppressed so that the SNR is relatively low. At these low elevation angles the direct signal and multipath signal interfere with each other. The oscillation interference pattern frequency is primarily a function of the differences in the path length. As the satellite arises the propagation of the signal which is received directly increases because of antenna's gain pattern. Approximately one hour is taken by the satellite to go from 5° to 30° . Below the elevation of 5° , it is not possible to capture the SNR data as the receiver does not accept the data from the low angle [4].

1.3 Neutron Thermalisation Technique

Since the neutrons are uncharged particle there is no effect of electromagnetic field on them. So the neutrons which are emitted from a source travel in a straight line toward the nucleus of an atom. When the neutron colloids with the hydrogen nucleus, the

neutron loses all of its energy and get thermalized. The collision results in a change in the direction of the neutron. Thus if we count the number of thermalized neutrons near the source within a period of time, the number of hydrogen nuclei in solids can be estimated which can be used for the derivation of the soil moisture content. Since the use of radioactive material is very harmful this device is less attractive for determining the moisture content in the soil [3].

1.4 Gamma Attenuation Technique

This technique can be used for determining the soil moisture with resolution below depth of 25 mm or less. The principle of this technique is based on the fact that absorption and scattering of the gamma rays have a relation with the density of matter in their path. The increase or decrease in the moisture of the soil changes the saturated density of soil whereas the specific gravity remains constant. The gamma transition technique is used for measuring the changes in saturated density and then the moisture content of the soil is determined by this density change. This technique can be used for automatic measurement and recording. The disadvantage of this technique is that the gamma rays used are very dangerous and also the cost of field operation is relatively high to use [4].

1.5 Remote Sensing

In this technique we get the information of the object without touching it physically. The soil moisture measurements are made using the instruments which are not connected directly to the soil. The changes in the moisture of soil are inferred through the influence of the soil in the potential field like magnetic field, electric field, and gravitation field. This method of remote sensing has distinct advantages and challenges as compared to ground-based techniques. In the current scenario, there are three major remote sensing methods which are used for the measurement of soil moisture. Out of these three methods, the first two methods consider either the electromagnetic radiation which is naturally emitted by the target or the radiation which are scattered by the target after it has been illuminated with a radiation from the known source. The third method detects the soil moisture through the changes in the gravity potential field which itself related to the changes in the density of the soil. The main advantage of this technique is that this method has the capacity of measuring the large areas with the single instrument and this technique can also be used in those isolated areas where in situ method cannot be feasible [5–10].

1.6 Global Navigation Satellite System (GNSS)

The GNSS reflectometry is a form of bistatic radar. Other systems for example those which are used for approaching harbor and monitoring airspace and for forecasting the weather is the form of monostatic radar in which both the radar transmitter and receiver are combined together at the same site. But in case of bistatic systems, the transmitter and receiver section are separated from each other by some considerable distance. Such type of system can be used for studying certain weather phenomena and some military operations where simply the LOS (line of site) reflections from the target are insufficient or inadequate. The concept of the bistatic radar system can be extended to the signals from the satellites. The Signals which are transmitted from the satellite get scattered by the surface of the earth thus if we detect these signals by some separate passive receiver it could provide us with some information about the surface.

The earthbound signals are indiscriminate and get transmitted by the GPS transmitting satellites and the receiving antennas passively intercept the incoming signals at the ground. The signal which gets reflected from the ground represents a source of noise that the geodesists try to suppress. The role of the GPS receiver is to record the information of noise in the environment as a signal to noise ratio interferograms. The data is specific to the track of the satellite such that each one of the track has its own SNR data time series on a given day. The received SNR data is then recorded as the function of time. This time is itself a function of the elevation angle of the satellite with respect to the horizon [5–11].

2 Conclusion

We have discussed some methods for determining the moisture content of the soil. In all the methods discussed above we can say that the GNSS methods are considered suitable for determining the moisture content of the soil since it can be used over the large experimental area, not like point measurement and it is less laborious as that of thermogravimetric method thus the GNSS technique provides good options for the determination of soil moisture content.

References

1. Liao, F., Porporato, A., Ridolfi, L., Rodriguez-Iturbe, I.: Plants in water controlled ecosystems: active role in hydrological processes and response to water stress. *Adv. Water Resour.* **24**, 707–723 (2001)
2. Robinson, D., Campbell, C., Hopmans, J., Hornbuckle, B., Jones, S.B., Knight, R., Ogden, F., Selker, J., Wendroth, O.: Soil moisture measurement for ecological and hydrological watershed-scale observatories: a review. *Vadose Zone J.* **7**(1), 358–389 (2008)

3. Su, S.L., Singh, D., Baghini, M.S.: A critical review of soil moisture measurement. *Measurement* **54**, 92–105 (2014)
4. Zazueta, F.S., Xin, J.: Soil moisture sensors. *Soil Sci.* **73**, 391–401 (1994)
5. Chew, C., Small, E.E., Larson, K.M.: An algorithm for soil moisture estimation using GPS-interferometric reflectometry for bare and vegetated soil. *GPS Solut.* 1–13. <https://doi.org/10.1007/s10291-015-0462-4> (2015)
6. Egido, A., Caparrini, M., Ruffini, G., Paloscia, S., Santi, E., Guerriero, L., Pierdicca, N., Flouri, N.: Global Navigation Satellite Systems Reflectometry as a Remote Sensing Tool for Agriculture. *Remote Sens.* **4**, 2356–2372 (2012)
7. Larson, K.M., Braun, J.J., Small, E.E., Larson, K.M., Braun, J.J., Small, E.E., Zavorotny, V.U., Gutmann, E.D., Bilich, A.L.: GPS Multipath and its relation to near-surface soil moisture content. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **3**, 91–99 (2010)
8. Larson, K., Braun, J., Small, E., Zavorotny, V., Gutmann, E., Bilich, A.: GPS multipath and its relation to near-surface soil moisture content. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **3**, 91–99 (2010)
9. Chew, C.C., Small, E.E., Larson, K.M., Zavorotny, V.U.: Effects of near-surface soil moisture on GPS SNR data: development of a retrieval algorithm for soil moisture. *IEEE Trans. Geosci. Remote Sens.* **52**, 537–543 (2014)
10. Njoku, E., Lakshmi, V., O'Neill, P.: Soil moisture field experiment special issue. *Remote Sens. Environ.* **92**, 425–426 (2004)
11. Hofmann-Wellenhof, B., Lichtenegger, H., Wasle, E.: GNSS Global Navigation Satellite Systems: GPS, GLONASS, Galileo & more. Springer, Wien, New York (2008). ISBN: 978-3-211-73012-6

Monostatic Radar Based Ultra-Wideband Microwave Imaging System Featuring a Miniature Fork Shaped Microstrip Patch Antenna with a Reduced DGS for Early Breast Tumor Detection



Arashpreet Kaur and Amanpreet Kaur

Abstract This paper presents the usage of a fork-shaped microstrip patch antenna for early breast tumor diagnosis based on significant contrast in dielectric properties between healthy and malignant tissue. The proposed antenna is designed with a 50Ω microstrip feed line and has compact dimensions of $24 \times 28 \times 1.64 \text{ mm}^3$ in CST MWS V'18. The antenna structure has a slotted rectangular patch with a reduced ground to achieve high gain, miniaturization, and UWB characteristics. A 3D spherical breast phantom is modeled with 3 mm tumor radius and simulated using CST MWS V'18 with different dielectric properties of skin, fatty tissue, and tumor. The breast phantom is oriented parallel to the broadside radiating surface of the antenna and the backscattered signals are recorded when the phantom is illuminated by the microwave signals for with and without tumor. The simulated results show that more reflections, lesser specific absorption rate and more conduction current density is obtained in the presence of tumor as compared to a nonmalignant case thereby making the detection of the tumor in the breast feasible.

Keywords Slotted MSA · Defected ground structure (DGS) · Ultra-wide band (UWB) · Breast tumor · Dielectric properties · Specific absorption rate(SAR) · Computer simulation tool microwave studio software (CST MWS'18)

1 Introduction

In the current world, breast cancer is the most frequently diagnosed cancer type and a second leading cause of female mortality in the USA. According to the Global Cancer Statistics in 2012, 1.7 million new cases of breast cancer were diagnosed and 521,900 women died of this disease worldwide [1]. Breast cancer occurs when

A. Kaur (✉) · A. Kaur
Antenna Research and Testing Laboratory, Department of Electronics and Communication Engineering, Thapar University, Patiala 147004, India
e-mail: sohi_arsh@gmail.com

the abnormal cells continue to grow uncontrollably even when the new cells are not required. The tumor may be benign or malignant in nature. Experts suggest that detection of breast cancer at an early stage is the best way for preventing the serious toll of this disease [2].

Several techniques are available for breast cancer detection such as X-ray mammography [3], Ultrasound [4], and Magnetic Resonance Imaging [5]. Microwave imaging is an effective method for early diagnosis and screening of breast cancer due to its innate advantages of nonionizing radiations, more safety, low cost, and sensitivity to tumors [6]. The working principle of microwave imaging is based on the concept of significant contrast in the dielectric properties (permittivity and conductivity) between the malignant and healthy breast tissues when the breast structure is illuminated by the electromagnetic waves at microwave frequencies (300 MHz to 30 GHz) [7].

The choice of antenna plays an essential role in UWB microwave imaging systems. Generally, low profile antennas are preferred which can exhibit large operational bandwidth in the range of 1–10 GHz for better penetration depth and spatial resolution [8]. Microstrip patch antenna is the preferred choice due its advantages of small size, light weight, low manufacturing cost, and conformity to both planar and nonplanar surfaces. But the microstrip patch antennas suffer from the problem of narrow impedance bandwidth and low gain [9]. Different techniques have been reported to overcome the shortcomings suffered by microstrip patch antenna such as stacked patch [10], use of DGS and fractal geometry [11], slotted geometry [12], etc.

Based on the literature survey carried out in the field of MSA, the prime objective of this research is to achieve miniaturization, high gain, and ultra-wideband characteristics from a slotted MSA with a reduced DGS for diagnosis of malignant tumors by microwave imaging system. The designed antenna covers an ultra-wideband from 3.71 to 11.48 GHz with bandwidth of 7.77 GHz and shows a peak gain of 5.82 dB at 11 GHz frequency. A spherical tumor of radius 3 mm is designed and embedded inside a 3D spherical breast phantom of radius 30 mm. The antenna is placed at a distance of 10 mm away from the skin. All the designing and simulations are implemented using Computer Simulation Tool Microwave Studio software (CST MWS'18). Active radar-based microwave imaging is employed which illuminates the breast structure by radiating a short pulse signal by a single UWB antenna and records any backscattered signal by the same antenna to analyze the difference between with and without tumor cases [7].

2 Antenna Design and Specifications

In order to achieve a UWB from a conventional slotted microstrip patch antenna, tuning stubs and a DGS are incorporated in the proposed design. The antenna prototype is photo-etched on an FR4 substrate with a dielectric constant of 4.4 and dielectric loss tangent value of 0.0024. The substrate dimensions for this microstrip line fed

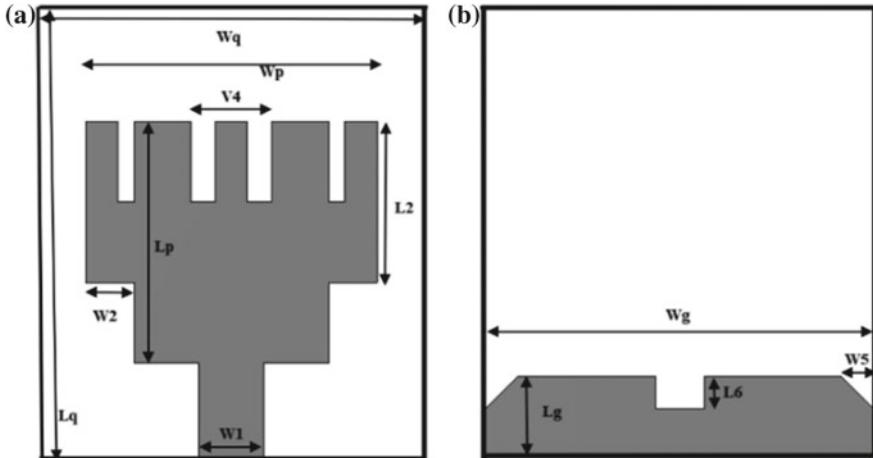


Fig. 1 **a** Top view of proposed antenna **b** Ground layer of proposed antenna with DGS

MSA are $24 \times 28 \times 1.57$ mm³. Figure 1a illustrates the top view of the proposed antenna design which shows an upturned U-slotted rectangular patch joined to two L-shaped tuning stubs (at the edges of the patch) which acts as radiators. The patch is fed along the centerline of symmetry using a $50\ \Omega$ microstrip feed line. Dimensions of the feed line are optimized to provide a good impedance matching between the generator impedance and input impedance of the radiating patch. The optimized dimensions of the proposed antenna with all parameters labeled in Fig. 1a and b are specified in Table 1. The design process starts with selecting a rectangular patch antenna with optimized dimensions. The designing is done using the transmission line model equations mentioned in CA Balanis [9]. An upturned U-shaped slot is truncated from the upper edge of the radiating patch based upon the current distributions observed on the patch which results in multiband behavior. In order to cover a UWB range, two L-shaped tuning stubs are joined on each side of the slotted rectangular patch which increase the electrical length of the patch and result in bandwidth enhancement.

Figure 1b demonstrates the bottom view of proposed antenna with a reduced ground structure and a rectangular slot etched out from it to improve the impedance matching antenna characteristics. In order to further enhance the antenna bandwidth, two right-angled isosceles triangular notches are notched from the left and right corners of the reduced DGS. Assuming perfect magnetic side walls, the side of the right-angled isosceles triangular notch for dominant mode TE₁₀ is calculated as follows:

$$A = \frac{c}{2f_r \sqrt{\epsilon_r}} \quad (1)$$

Table 1 Proposed antenna design parameter

Parameter	Dimension (mm)
Lq	28
Wq	24
Lp	15
Wp	12
L2	10
W1	4
W2	3
V4	5
Lg	5
Wg	24
W5	2
L6	2

where c , f_r , ϵ_r , A are speed of light, resonant frequency, dielectric constant, and side of right-angled isosceles triangular notch, respectively [15].

3 Simulation Results and Analysis

To prove the suitability of the proposed antenna for microwave imaging systems, its performance is studied in terms of return loss, bandwidth, and gain using CST MWS'18.

3.1 Return Loss and Bandwidth

Figure 2 shows a graph of S_{11} (dB) values which varies with respect to the frequency. The frequencies with S_{11} (dB) value of -10 dB and below occupies the bandwidth of 7.77 GHz from 3.71 GHz to 11.48 GHz. The peak value of S_{11} (dB) is -57.71 dB at 4.55 GHz resonant frequency. So the antenna shows good impedance matching and minimal power loss.

3.2 Gain

Figure 3 shows broadband gain of the designed antenna over the entire operational band. For the frequency range from 8.68 GHz to 11.48 GHz, the gain is greater than 4 dB and a peak gain of 5.82 dB is achieved at the frequency of 11 GHz. The elevation and azimuth plane pattern of antenna's gain at 11 GHz are also shown in Fig. 4(a–b).

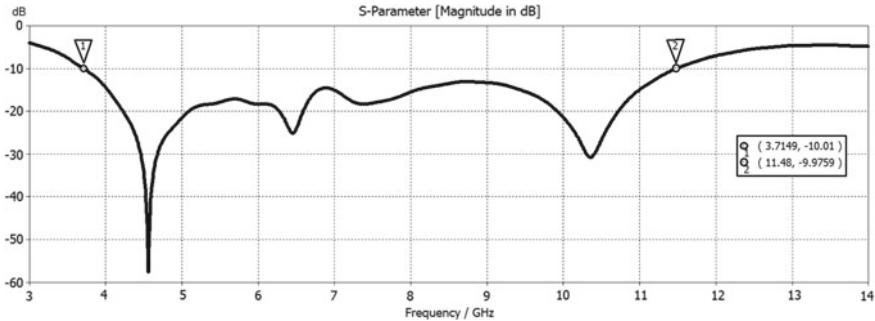


Fig. 2 Simulated S_{11} dB plot with respect to frequency

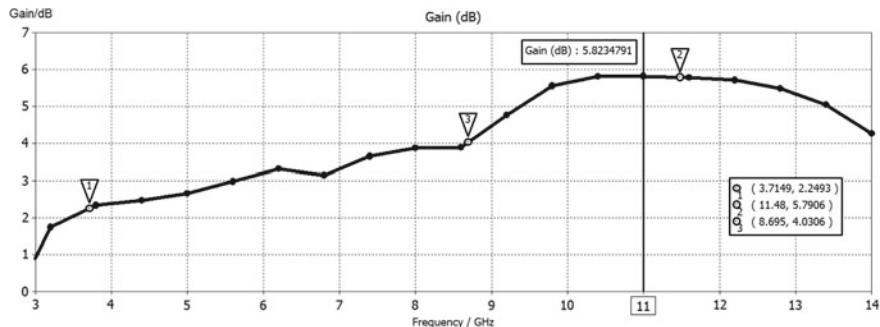


Fig. 3 Broadband gain plot with respect to frequency

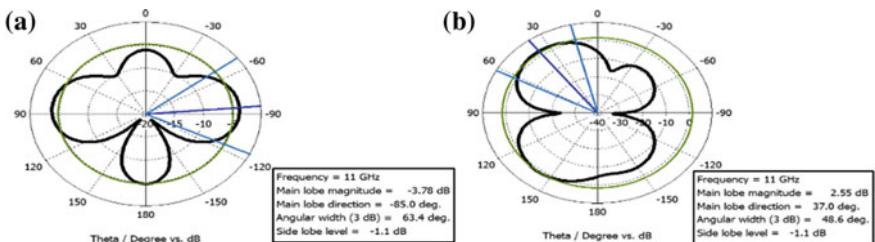


Fig. 4 Polar plot of the proposed antenna at 11 GHz for **a** Elevation plane **b** Azimuth plane

4 Spherical Breast Phantom Model for Tumor Detection

Figure 5 shows a 3D spherical breast structure which is a three layered model where the outer skin is 2 mm thick and other two layers are inner fatty tissue and a spherical tumor with 58 mm and 6 mm diameter, respectively. Table 2 summarizes different electrical properties of all three breast tissues [7]. To diagnose the presence of tumor, the simulations are carried out by placing the breast phantom at 10 mm distance along the z-axis from the proposed UWB antenna. Microwave imaging concept is based

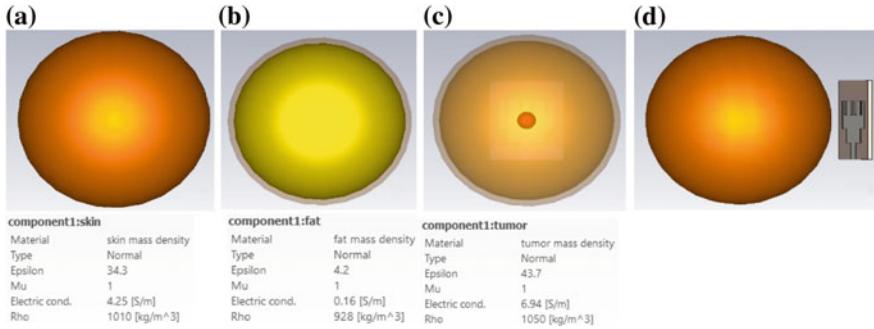


Fig. 5 Spherical breast phantom model **a** Skin layer of 60 mm diameter **b** Fatty layer of 58 mm diameter **c** Tumor of 3 mm radius **d** Phantom placed at 10 mm distance from antenna

Table 2 Dielectric properties of breast tissues at 4 GHz frequency [7]

Tissue type	Permittivity	Conductivity (S/m)	Mass density (Kg/m ³)
Skin layer	34.3	4.25	1010
Fatty tissue	4.2	0.16	928
Tumor	43.7	6.94	1050

on the tissue dependent microwave scattering and absorption which measures the significant contrast in dielectric properties of skin, fatty tissue, and tumor. Due to the low water content of fatty tissue, it is characterized by low values of permittivity and conductivity whereas the high water content tumor is featured with high values of permittivity and conductivity which causes larger microwave scattering as compared to normal breast tissue [8].

5 Simulation Results and Discussions for Breast Tumor Detection

To detect the presence of tumor in the breast phantom, simulations are carried out for Return loss, SAR, and current density in CST MWS'18.

5.1 Return Loss Characteristics

Figure 6 shows the comparison of return loss characteristics for without and with tumor cases by placing the antenna at a 10 mm distance from the phantom. So in the frequency band from 3.7537–8 GHz with peak return losses of -26.38 dB (without tumor) and -21.619 dB (with tumor) at 4.35 GHz resonant frequency shows

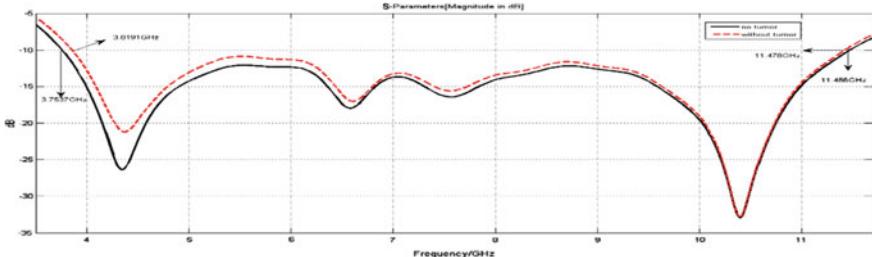


Fig. 6 Comparison of Simulated S_{11} dB plot with respect to frequency for breast phantom model without and with tumor cases

more positive reflection coefficient as the plot shifts upwards (dotted red line) in the presence of tumor as compared to the normal skin tissue (solid black line). So the presence of 3 mm (radius) tumor is easily noticeable. It is because in the presence of tumor the received reflected signal is more which results in poor impedance matching and poor return loss characteristics as compared to the nonmalignant case. The frequency band from 8–11.47 GHz shows almost insignificant variations in terms of peak return loss magnitude that is -32.77 dB (without tumor) and -33.14 dB (with tumor) at 10.41 GHz resonant frequency. So the existence of tumor is not noticeable in this band.

5.2 Specific Absorption Rate (SAR) and Conduction Current Density

SAR is the measure of the rate at which RF energy is absorbed by the living tissues when exposed to the electromagnetic field. It is the power absorbed per tissue mass averaged over whole body or a small volume. SAR can be evaluated as follows:

$$\text{SAR} = \frac{\sigma |E|^2}{\rho} \quad (2)$$

where E is rms electric field (V/m), σ is tissue conductivity (S/m), and ρ is tissue mass density (Kg/m^3) [2]. Figure 7(a–b) shows that the peak SAR value on skin for 1 g tissue is less (12.7 W/Kg) in the presence of tumor as compared to nonmalignant case (20.4 W/Kg) at 4.35 GHz peak frequency due to more power loss from the affected area. Figure 7(c – d) shows that peak current density is more (721.1 A/m^2) in the presence of tumor as compared to nonmalignant case (800.9 A/m^2) at 4.35 GHz peak frequency.

Since the breast phantom in the presence of tumor results in lesser SAR(1 g), more current density, and more reflections with 4.761 dB variation in the return loss magnitude at 4.35 GHz resonant frequency as compared to a nonmalignant case, the

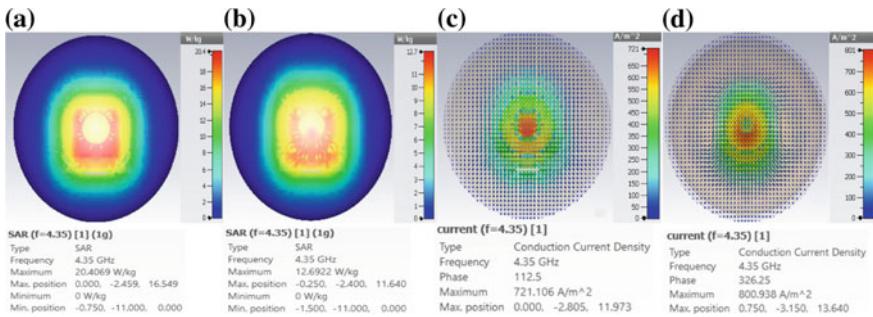
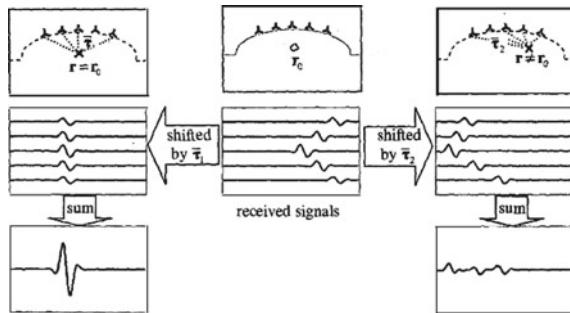


Fig. 7 Comparison at 4.35 GHz frequency for breast phantom model for **a** SAR without tumor **b** SAR with tumor **c** Current density without tumor **d** Current density with tumor

Table 3 Comparison of the stimulated results in terms of return loss, SAR, and current density in the frequency band from 3.7537–8 GHz at resonant frequency of 4.35 GHz for with and without tumor

Parameters	Without tumor	With tumor
Return loss (dB)	-26.38	4.25
SAR (W/Kg)	20.4069	0.16
Current density (A/m ²)	721.106	6.94

Fig. 8 Demonstration of delay and sum beamforming algorithm [14]



presence of the tumor is clearly detectable. The comparison of simulated results for with and without tumor case are illustrated in Table 3. Further, an array configuration can be considered and arranged around the phantom in order to increase the sensitivity in determining the presence of a tumor. Time delayed version of S_{11} (dB) can be used in time shifting and adding algorithms such as delay and sum beamforming in order to develop an image of the illuminated breast phantom with a tumor as shown in Fig. 8.

6 Conclusion

A compact fork-shaped microstrip patch antenna with reduced DGS has been successfully prototyped and simulated in this article. The proposed UWB antenna covers a UWB range from 3.71 to 11.48 GHz with a bandwidth of 7.77 GHz and offers an average overall gain of 4.03 dB for the entire frequency band. This compact geometry and required bandwidth with a good average gain helps in successfully using this antenna for microwave imaging systems for effectively detecting the malignant tumors growing inside the breast. A 3D spherical breast model with tumor radius of 3 mm was designed and the proposed antenna was placed at 10 mm distance from the phantom. It is observed that in the presence of tumor the received reflected signal is more which results in poor impedance matching, less peak SAR value, and more peak current density as compared to the nonmalignant case. In future, genetic algorithms can be used for increasing the sensitivity in identifying the location of tumors.

Acknowledgments The authors are thankful to University Grants Commission, New Delhi, India has provided a Fund grant of 15.8 lakhs under the major project scheme for setting up the antenna testing facility at Thapar University, Patiala.

References

1. Torre, L.A., Bray, F., Siegel, R.L., et al.: A Global cancer statistics. *Cancer J. Clin.* **65**, 87–108 (2015)
2. Karli, R., Ammor, H.: Early detection of breast tumors using UWB microstrip antenna imaging. *Microw. Opt. Technol. Lett.* **58**, 2101–2106 (2016)
3. Mettler, F.A., Upton, A.C., Kelsey, C.A., et al.: Benefits versus risks from mammography: a critical reassessment. *Cancer* **77**, 903–909 (1996)
4. Gordan, P.B., Goldenberg, S.L.: Malignant breast masses detected only by ultrasound. A retrospective review. *Cancer* **76**, 626–630 (1995)
5. Haran, E.F., Grobgeld, D., et al.: Critical role of spatial resolution in dynamic contrast-enhanced breast MRI. *J. Magn. Reson. Imaging* **13**, 862–867 (2001)
6. Nikolova, N.K.: Microwave imaging for breast cancer. *IEEE Microw. Mag.* **12**, 78–94 (2011)
7. Sill, J.M., Fear, E.C.: Tissue sensing adaptive radar for breast cancer detection-experimental investigation of simple tumor models. *IEEE Trans. Microw. Theory Tech.* **53**, 3312–3319 (2005)
8. Moore, S.K.: Better breast cancer detection. *IEEE Spectr.* **38**, 50–54 (2001)
9. Balanis, C.A.: *Antenna Theory Analysis and Design*. Wiley, Hoboken (2005)
10. Kaur, A., Khanna, R.: Design and development of a stacked complementary microstrip antenna with a ‘ π ’-shaped DGS for UWB, WLAN, WiMAX and radio astronomy wireless applications. *Int. J. Microw. Wirel. Technol.* **2**, 1–10 (2017)
11. Kaur, A., Khanna, R., Kartikeyan, M.V.: A stacked sierpinski gasket fractal antenna with defected ground structure for UWB/WLAN/Radio astronomy/STM link applications. *Microw. Opt. Technol. Lett.* **57**, 2786–2792 (2015)
12. Vuong, T.P., Ghiootto, A., et al.: Design and characteristics of a small U-slotted planar antenna for IR-UWB. *Microw. Opt. Technol. Lett.* **49**, 1727–1731 (2007)

13. Maity, S., Gupta, B.: Accurate resonant frequency of isosceles right-angled triangular patch antenna. *Microw. Opt. Technol. Lett.* **55**, 1306–1308 (2013)
14. Li, X., Bond, E., Van Veen, B., Hagness, S.: An overview of ultra-wideband microwave imaging via space-time beam forming for early-stage breast cancer detection. *IEEE Antennas Propag. Mag.* **47**, 19–34 (2005)

An Enhanced Authentication Technique to Mitigate the Online Transaction Fraud



Vipin Khattri, Sandeep Kumar Nayak and Deepak Kumar Singh

Abstract Online transactions provide benefits to the customer and financial organization in terms of reducing operating cost, time, efforts, papers, and increasing the comfort and ease. Beside these benefits online transaction has a side effect like forged online transaction. This side effect results in loss of money. This forged online transaction is performed by fraudsters who are well equipped with the dynamic novel idea to steal an amount of money from the customer through online transaction. Although security measures are already implemented this forged online transactions are increasing every year. This is happening due to fraudsters creating novel ideas continuously to perform forged online transaction. Therefore, security measures need improvements continuously on a regular basis. The key aspiration of this paper is to create an improved authentication technique to prevent forged online transaction. This study produces an enhanced authentication using a mobile application to mitigate online transaction fraud.

Keywords Forged online transaction · Online transaction · Multi-factor authentication · One time password · Authentication · Authorization

V. Khattri · S. K. Nayak (✉)

Department of Computer Application, Integral University, Dasauli, Basha Kursi Road, Lucknow 226026, India

e-mail: nayak.kr.sandeep@gmail.com

V. Khattri

e-mail: vipinkhattri@gmail.com

D. K. Singh

Jaipuria Institute of Management, Vineet Khand, Gomti Nagar, Lucknow 226010, India

e-mail: deepak.iiita@gmail.com

1 Introduction

In the vision of reducing paperwork, time of transaction, manpower, resources, mass of crowd in a financial organization and increased security, comfort, digital transactions were established. It indicates that instead of going to a financial organization to take money, user can purchase products online, withdraw amount from an ATM and purchase the product from a shop using point of sale (POS). This convenience catches the attention of fraudsters and encourages them to attempt forged online transaction. The punishment of forged online transaction is money loss, effort loss, time loss, resource loss, and majorly infringement of trust [1]. To cope with forged online transaction, number of authentication [2–4] and authorization [5–15] techniques have been applied but forged online transaction is on top and increasing at an alarming rate. To keep in mind the aforesaid fact, authors put forward a projected enhanced authentication technique to mitigate forged online transaction. This study is prepared by authors and divided into five parts. The second part describes the variety of forged online transactions and techniques for mitigating forged online transaction. The third part sets up an enhanced authentication technique using mobile application developed by authors. The fourth part of the study describes the case that is associated with forged online transactions and influence of projected enhanced authentication technique in the case of forged online transaction. The fifth part of the paper, draws a conclusion and works for the future of the study.

2 Frauds of Online Transaction and Techniques for Mitigating Forged Online Transaction

2.1 *Frauds of Online Transaction*

Forged Online Transaction with ATM. To perform forged online transaction with ATM, two devices are used. The first device is a skimmer device and the second device is a forged keypad or hidden camera. This skimmer device is used to copy the facts or details of the transaction card and the forged keypad is used to trace the PIN of transaction card while punching on the keypad. The skimmer device is fixed on the slot of swap card. A forged keypad is fixed on the keypad of ATM. When a customer performs a transaction, all the details of the transaction card and PIN are recorded. Later the fraudster obtains these details from the skimmer device and forged keypad. These details are used by the fraudster to perform forged online transactions.

Forged Online Transaction with Phone. Fraudsters call the customer as a victim through phone and give some mischievous information such as security verification or corruption of data or system maintenance or penalty of card or blocking of card and request for verification of transaction card. Simultaneously fraudsters conduct a forged online transaction. After that the fraudsters take all the information including

OTP from the victim and perform forged online transaction. Fraudsters attempt this fraud in a very clean way that the user is not able to actually understand or think about the forged online transaction.

Forged Online Transaction with Cloned Transaction Card. With increase of digital transaction, customer can buy the product and pay with transaction card using point of sale device. Therefore fraudsters make a target of those shops which are equipped with POS device. In these targeted shops, fraudsters place a skimmer device with the help of staff members. When the customer swaps the card on the POS device, the staff member swaps the card on hidden skimmer device also and gazes the customer's activity while entering the PIN number. This information transfers to the fraudsters. Fraudsters use this information to make a duplicate card by integrating the actual card details. This cloned card is used for forged online transaction with PIN.

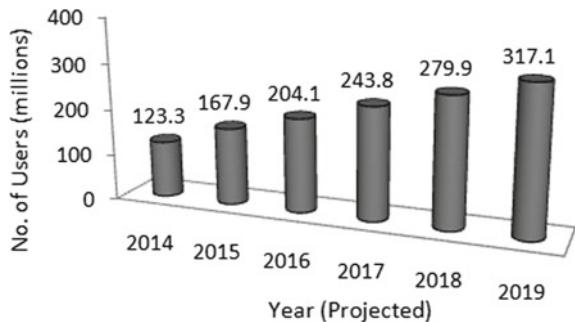
2.2 Existing Techniques of Authentication & Authorization

The variety of techniques is applied to mitigate forged online transaction at different levels of security. These levels are explained in the subsequent paragraphs.

Level 1-Authentication. At the first level of authentication, customer inputs the details of transaction card for authenticating initially [16–18]. An authentication using multi-factor authentication is put into action to formulate online transaction more safe and secure. The reason for applying for authentication using multi-factor is to forbid an unauthorized user. When authentication is using a single factor only then it is risky for online transactions because details of the single factor are static in nature and this detail is easy to steal. To compensate for this risk, multi-factor authentication is used. The details of second factor is dynamic, i.e., information of second factor is changed on every new transaction.

Level 2-Authorization. After verifying the details at the first level of authentication, the next level is authorization. The authorization confirms for the validity of a legitimate online transaction. To confirm the validity, various authorization techniques are implemented. These techniques work on the basis of human behavior for conducting online transaction. The variety of techniques are Data Mining [5] which works on the pattern of transaction, Hidden Markov Model [6, 7] which works on hidden states of transaction, Artificial Neural Network [8–11] which works on parameters of online transaction as input node, Artificial Immune System [12, 13] works on the basis of immune system and Genetic algorithm [14, 15] which works on finding the forged online transaction using biological evolution and many more.

Fig. 1 Users of mobile phone [19]



3 Projected Enhanced Authentication Technique

Customer performs an online transaction using transaction card and PIN/OTP number. Fraudster uses the same information (stolen) to perform forged online transaction without the knowledge of the customer. This means that the customer knows about the forged online transaction after completing it. Therefore some mechanism is required so that customer could know first about the online transaction and after confirmation the online transaction could proceed. To accomplish the same, authors introduce a third factor of multi-factor authentication. The third factor is mobile application. Phone is an integral part of the life of a human. In the digital world, a cellular phone is required and it is increasing at an exponential rate. From the graph (as shown in the Fig. 1) the users of cellular phone are increasing every year approximately at a growing rate of 20–30%. Fifty percent users are using their phones for online transaction [19]. In the projected enhanced authentication technique, there is a requirement for opening mobile application before initiating the session of online transaction. For completing the transaction, there is a requirement of OTP verification from the mobile application. These two things enhance the authentication and mitigate the forged online transaction.

3.1 Key Components of Authentication

Cellular Phone. This proposed technique works with a cellular phone that should be armed with mobile application for transaction. Online transaction cannot be initiated without using a cellular phone.

Mobile Application. Before applying the proposed technique, there is a requirement of mobile application which is designed by the financial organization and installed in the cellular phone with registered phone number of the customer. This application is required for initiating an online transaction. The utility of this application has two folds. First is to initiate the online transaction and second is to authorize the transaction by identifying OTP sent by the financial organization during the online

transaction. During the activation of this mobile application, no phone calls can be received or dialed by the user. This restricts the phone fraud.

Operational Device for Transaction. There is a requirement of a device through which the user can execute an online transaction. This device can be a personal computer, point of sale, or automatic teller machine.

Transaction Card. The transaction card is required to carry out an online transaction. This transaction card should be present with the authorized user. This transaction card is used for entering the details through the website or swapping card on ATM or POS device for carrying out the online transaction.

3.2 Procedure for Enhanced Authentication Technique

The core intention for developing the proposed technique is that after stealing all the information of transaction card including PIN number or transaction card with PIN number but in any circumstances fraudster will not be able to perform forged online transaction. To implement this intention, a mobile application for transaction plays an important role for performing legitimate online transaction. This can be possible through the mobile application which is used for the online transaction. Online transaction will be initiated after opening the mobile application. The procedure as shown in Fig. 2 along with flow chart as shown in Fig. 3 is explained given below:

Procedural Step-1-Initiate transaction by starting the mobile application for conducting the online transaction using some security code.

[Restriction-1: This mobile application should be set up in cellular phone with the linked phone number of the account number.

Restriction-2: During the activation of this mobile application, no phone calls can be received or dialed by user.

Fig. 2 Enhanced authentication procedure

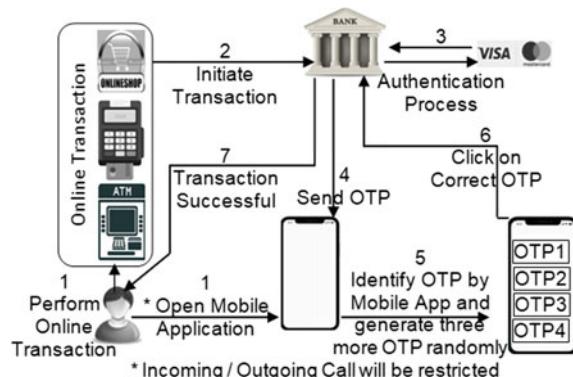
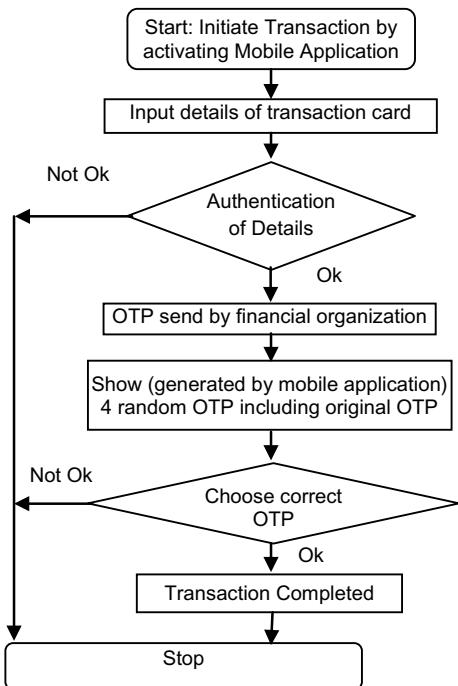


Fig. 3 Flowchart of an enhanced authentication technique



Restriction-3: Mobile application should be opened for starting the online transaction.]

Procedural Step-2-Initiate transaction by providing the credentials through transaction device.

Procedural Step-3-Authentication of details of the user by the financial organization through the transaction card's authority (e.g., VISA).

If authentication details are true

then

An OTP is generated by the financial organization and delivered to the registered phone number and follow step 4

else

Transaction aborted with denial message and follow step 8.

Endif

Procedural Step-4-An OTP is sent by the financial organization and is received by the registered cell number.

Procedural Step-5-The mobile application shows four different OTPs including original OTP received from the financial organization.

[The original OTP is recognized by mobile application as it is received by the phone and randomly 3 more OTPs are generated by the mobile application.]

Procedural Step-6-Select OTP which is original/correct.

[Restriction-4: Time limit for selection is 45 s.]

If Correct OTP is chosen

then

Completion of transaction

Else

Transaction aborted

Endif

Procedural Step 7-End of the Process

The projected authentication technique has a positive influence on different forged online transaction even though it is initiated by point of sale or ATM or personal computer or any other device. The analysis of impact is also shown in subsequent paragraphs.

Sneaked Transaction Card. After applying the projected authentication technique, it is observed that after sneaking the transaction card the fraudster cannot initiate forged online transaction because for starting the session of transaction using mobile application is required and which can't be possible for the fraudster and finally no forged transaction is initiated.

Forged Online Transaction using Phone. Author observed that after applying projected authentication technique, it is found that forged online transaction cannot be performed. Since the fraudster is unable to open the application for transaction and fraudster cannot make a call to the victim to perform forged online transaction.

The vital role of the projected enhanced authentication technique is to perform legal online transaction and helps for differentiating the genuine and forged online transaction. The main factor of projected technique is the mobile application with OTP verification on mobile application and this prevents the fraudster to operate forged online transaction.

4 Impact of Proposed Enhanced Authentication

4.1 Case of Forged Online Transaction Using Phone

This case was done with someone (victim) who lived in Uttar Pradesh, India. A person made a call to the victim and introduced himself as agent RRR (name changed) of a financial organization. Victim was notified by the fraudster that the transaction card was blocked and a penalty of Rs. 5000/- was enforced. The fraudster suggested

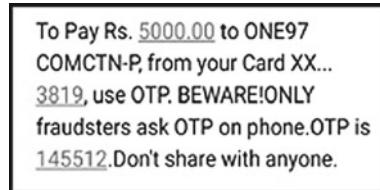


Fig. 4 Message of One Time Password (OTP)

verifying the details of the transaction card for unblocking the card and removing the penalty. The victim told all the details of the transaction card one by one as asked by the fraudster. In reality, the fraudster was performing forged online transaction at his end simultaneously. At last the fraudster asked for an OTP (Fig. 4) which was received by the victim and the victim gave the same. As the fraudster put an OTP for transaction the forged online transaction was completed and Rs. 5000/- amount should be withdrawn. Fortunately money in account was less than Rs. 1000/-. Therefore no amount was deducted but rejection charge due to lack of amount in account was enforced.

4.2 Analysis of Impact of Proposed Enhanced Authentication

Forged online transaction was committed in the previously depicted case. If projected authentication is applied on it then forged online transaction will not be performed. For initiating online transaction, mobile application is necessary and authentication (OTP verification) of online transaction which is impossible for the fraudster. Therefore no forged online transaction will be committed.

5 Conclusion with Future Work

The trouble of forged online transaction is not new in the digital transaction. The popularity of digital transaction is increased due to its simplicity. In this study, to control the forged online transactions, the authors have analyzed the drawbacks of the system using different types of forged online transactions. To deal with the negative aspect, authors portrayed an enhanced authentication technique to mitigate forged online transactions. Author has also analyzed the impact of the technique on forged online transaction case. In the future, authors will take an effort to implement on the real system and will analyze the strength, validity, accuracy, and easiness of an enhanced authentication technique.

Acknowledgments We extend our gratitude toward the Integral University for acknowledging our research work and providing us with Manuscript Communication Number-IU/R&D/2018-MCN000407. We also extend the same toward the Shri Ramswaroop Memorial University for giving us financial support for our research work.

References

1. Bhatla, T.P., Prabhu, V., Dua, A.: Understanding credit card frauds. *Cards Bus. Rev.* **1**(6), (2003)
2. Kulat, A., Kulkarni, R., Bhagwat, N., Desai, K., Kulkarni, M.P.: Prevention of online transaction frauds using OTP generation based on dual layer security mechanism. *Int. Res. J. Eng. Technol. (IRJET)*, **3**(4), 1058–1060 (2016)
3. Jamdar, P.A., Bendale, S.K., Durgawale, D.A.: Secure mechanism for credit/debit card transaction fraud detection by using fingerprint authentication system. *Int. J. Comput. Sci. Netw. Secur. (IJCSNS)*, **16**(3), 35–38 (2016)
4. Khattri, V., Singh, D.K.: A Novel Distance Authentication Mechanism to Prevent the Online Transaction Fraud, pp. 157–169. *Advances in Fire and Process Safety*, Springer, Singapore (2018)
5. Singh, P., Singh, M.: Fraud detection by monitoring customer behavior and activities. *Int. J. Comput. Appl.* **111**(11), 23–32 (2015)
6. Prakash, A., Chandrasekar, C.: An optimized multiple semi-hidden markov model for credit card fraud detection. *Indian J. Sci. Technol.* **8**(2), 165–171 (2015)
7. Khan, M.Z., Pathan, J.D., Ahmed, A.H.E.: Credit card fraud detection system using hidden markov model and k-clustering. *Int. J. Adv. Res. Comput. Commun. Eng.* **3**(2), 5458–5461 (2014)
8. Bekirev, A.S., Klimov, V.V., Kuzin, M.V., Shchukin, B.A.: Payment card fraud detection using neural network committee and clustering. *Opt. Mem. Neural Netw.* **24**(3), 193–200 (2015)
9. Van Vlasselaer, V., Bravo, C., Caelen, O., Eliassi-Rad, T., Akoglu, L., Snoeck, M., Baesens, B.: APATE: A novel approach for automated credit card transaction fraud detection using network-based extensions. *Decis. Support. Syst.* **75**, 38–48 (2015)
10. Khan, A.U.S., Akhtar, N., Qureshi, M.N.: Real-time credit-card fraud detection using artificial neural network tuned by simulated annealing algorithm. In: *Proceedings of International Conference on Recent Trends in Information, Telecommunication and Computing*, pp. 113–121. (2014)
11. Behera, T.K., Panigrahi, S.: Credit card fraud detection: a hybrid approach using fuzzy clustering & neural network. In: *Second International Conference on Advances in Computing and Communication Engineering (ICACCE)*, pp. 494–499. (2015)
12. Halvaiee, N.S., Akbari, M.K.: A novel model for credit card fraud detection using Artificial Immune Systems. *Appl. Soft Comput.* **24**, 40–49 (2014)
13. Soltani, N., Akbari, M.K., Javan, M.S.: A new user-based model for credit card fraud detection based on artificial immune system. In: *16th IEEE CSI International Symposium on Artificial Intelligence and Signal Processing*, pp. 029–033. (2012)
14. Assis, C.A., Pereira, A., Pereira, M.A., Carrano, E.G.: A genetic programming approach for fraud detection in electronic transactions. In: *IEEE Symposium on Computational Intelligence in Cyber Security*, pp. 1–8. (2014)
15. RamaKalyani, K., UmaDevi, D.: Fraud detection of credit card payment system by genetic algorithm. *Int. J. Sci. & Eng. Res.* **3**(7), 1–6 (2012)
16. Kennedy, E., Millard, C.: Data security and multi-factor authentication: Analysis of requirements under EU law and in selected EU Member States. *Comput. Law & Secur. Rev.* **32**(1), 91–110 (2016)
17. Zhang, F., Kondoro, A., Muftic, S.: Location-based authentication and authorization using smart phones. In: *11th IEEE International Conference on Trust, Security and Privacy in Computing and Communications*, pp. 1285–1292. (2012)

18. Aloul, F. A., Zahidi, S., El-Hajj, W.: Two factor authentication using mobile phones. In: IEEE Conference on Computer Systems and Applications, pp. 641–644. (2009)
19. [www.eMarketer.com: https://apsalar.com/2016/03/mobile-and-smartphone-usage-statistics-for-india/](https://apsalar.com/2016/03/mobile-and-smartphone-usage-statistics-for-india/). Accessed on June 2018

A Study of Factors to Predict At-Risk Students Based on Machine Learning Techniques



Anu Marwaha and Anshu Singla

Abstract It is necessary, in education to identify the students who are facing problems in studies so that preventive actions can be taken to improve the students' performance. Early prediction of student performance is one of Educational Data Mining (EDM) application. In this paper, a set of factors like (academic, demographic, social, and behavior) and their influence on student performance have been studied for early prediction system. The authors have applied seven different machine learning (ML) models on the real data of students of Chitkara University, India. The study shows that to take preventive measures for students at risk not only current performance but the previous academic performance and demographical factors also play an important role. Ensemble model provides the most accurate results.

Keywords Learning analytics · Predictive modeling · Education data · Machine learning

1 Introduction

It is important to find students who are at risk as soon as possible during a course. With the limitation of a current academic early warning system, it is important to develop prediction models with high accuracy for at-risk students [1]. Predicting students' performance goal is to estimate the students' score from different aspects or behavioral factors of students. Apart from the previous schooling, prior academic

Please note that the LNCS Editorial assumes that all authors have used the western naming convention with given names preceding surnames. This determines the structure of the names in the running heads and the author index.

A. Marwaha (✉) · A. Singla

Chitkara University Institute of Engineering and Technology, Chitkara University, Chandigarh, India

e-mail: anumarwaha2010@gmail.com

A. Singla

e-mail: asheesingla@gmail.com

performance there exist a number of other factors, demographic, social, cultural, family socioeconomic status, psychological profile, which can affect the performance of students [2].

Several studies have been carried out by researchers to predict the students' academic performance in various tests and exams [3–11]. In this paper, for predicting students' performance and analyzing the effect of factors, authors compare seven different predictive ML techniques using various factors that influence students' performance.

The objectives of the present study are as follows:

1. To identify the most influential factors which are responsible for predicting the student at risk.
2. To identify which ML model is best to predict students' performance.
3. Is it possible to make early prediction at the start of a course or is it necessary to wait until mid-semester to obtain prediction with more accuracy?

The paper is arranged as follows: Sect. 2 gives a brief overview of the proposed model describing the techniques employed for data gathering, data processing, and learning. Section 3 describes student attribute and data set used in this work along with experimental settings followed by results discussion.

2 Proposed Model

This work proposes the model to identify the most influential factors that can predict students' performance in which not only students' pass or fail status is predicted but also how a student will perform (excellent, good, or average) also come under consideration. The main stages of the proposed model in the context of predicting the students' performance are shown in Fig. 1.

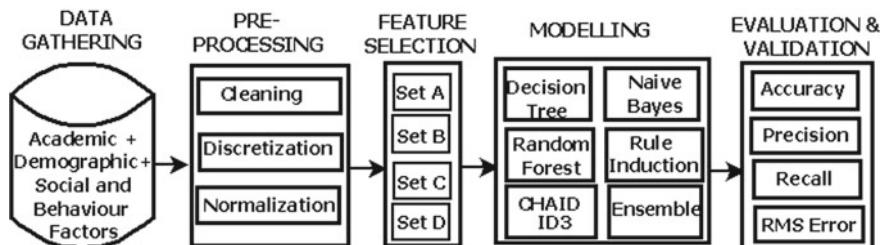


Fig. 1 Proposed model to predict the most influential factors of students at risk

Table 1 Factors considered to predict students at risk

Factors	Description
Academic	Defines the academic performance such as marks
Demographical	Define the characteristics of a person or a population such as race, age, and income
Social and behavioral	Define the lifestyle and habits of a person such as religion, family income, type, and so on

2.1 Data Gathering

Students performance is not only dependent on academic progress but on various other factors also. Table 1 shows the description of factors the authors have considered to predict students' performance based on literature [12]. Data Gathering is the most fundamental and essential stage of data analysis. The real data of 2560 students has been gathered from Chitkara University, India through online questionnaires.

2.2 Preprocessing Data

Data gathered from the different sources is loosely coupled. This results in redundant, out of range, incomplete, and inconsistent data [13]. With the increase in data, it may contain noisy and irrelevant information. Therefore, the gathered data goes through following a series of steps during the preprocessing phase.

2.2.1 Data Cleaning

Data discrepancy and outliers have been detected and various steps are followed to make anomalies free data. For instance, missing values in the 10th and 12th percentage have been detected and replaced by average marks.

2.2.2 Data Discretization

Continuous attribute values are divided into the range of attribute intervals on which classification techniques can apply. Certain standards for students' performance and parents' income have been followed to categorize them in different classes to apply ML techniques easily.

2.2.3 Data Normalization

Comparing attributes on different scales is not possible, so it is necessary to normalize the data in the same range on which they can be compared and prediction technique can be applied.

2.3 Feature Selection

The goal of feature selection is to reduce the dimension of the dataset and improve the accuracy of prediction [14, 15]. But evaluating every dimension and subset is practically infeasible. Therefore, on the basis of correlation highly correlated features are determined and only one of them is selected. To determine which factors are most influential, four different datasets combinations have been generated as described in Table 2.

2.4 Machine Learning Models

After the data preprocessing and feature selection, the preprocessed dataset can be used for applying classification models. Decision Tree (DT) [16], Iterative Dichotomiser 3 (ID3) [17], Chi-squared Automatic Interaction Detector (CHAID) [18], Naive Bayes Classifier (NBC), Rule Induction [19], Random Forest [20], Ensemble (EN) [21, 22], learning models were chosen to identify at-risk students. Among them DT, ID3, CHAID, and random forest all are various form of trees with different objective function and pruning level to make a decision tree more accurately.

Table 2 Factors studied to predict student at risk

Data set	Attributes
Dataset A	Academic demographic, social, and behavioral factors with prior semester performance
Dataset B	Academic demographic, social, and behavioral factors without prior semester performance
Dataset C	Demographic, social, and behavioral factors only
Dataset D	Academic factors only

2.5 Evaluation and Validation

For evaluation of the performance, four standard evaluation metrics (precision, recall, accuracy, and root mean square values (RMS)) are used [23]. The parameters are calculated from true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values calculated from the confusion matrix. Deviation between forecast (F) and observed (O) values are measured using RMS error. These performance evaluation parameters are defined as

$$\text{Accuracy} = \frac{TN + TP}{TN + FP + FN + TP} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N (F - O_i)^2} \quad (4)$$

3 Experiment and Results

For experimental purpose, the data of undergraduate students have been collected from Chitkara University, India from the year 2012 to 2016. Initially, records of about 2560 student were collected on the basis of 30 features which includes all factors as described in Table 2. Preprocessing and feature selection are applied to acquire the most pertinent attributes of students. After eliminating inconsistencies and duplications in the dataset, we considered 2156 student instances and 23 factors for experiments. Experiments have been carried out using 7.1 Rapid Miner tool.

Experiment 1: In order to study if the late prediction of student performance gives more accurate results by including prior semester results, ML models have been applied on dataset A. Table 3 shows the results of different evaluation parameters for different ML models.

Table 3 Results of various ML techniques using dataset A

Model	DT	ID3	CHAID	NBC	RI	RF	EN
Accuracy	77.42	58.06	70.97	74.19	74.19	74.42	83.33
Precision	84.00	78.95	85.71	80.77	78.57	79.31	82.76
Recall	87.50	62.50	75.00	87.50	91.67	95.83	100
RMS error	0.452	0.643	0.464	0.427	0.467	0.427	0.408

Table 4 Results of various ML techniques using dataset B

Model	DT	ID3	CHAID	NBC	RI	RF	EN
Accuracy	64.52	58.06	61.29	74.19	77.42	77.42	80.00
Precision	78.26	78.95	77.27	80.77	77.42	77.42	80.00
Recall	75.00	62.50	70.83	87.50	100	100	100.0
RMS error	0.585	0.644	0.610	0.478	0.460	0.449	0.398

Table 5 Results of various ML techniques using dataset C

Model	DT	ID3	CHAID	NBC	RI	RF	EN
Accuracy	70.97	58.06	64.52	67.74	77.42	77.42	80.00
Precision	77.42	74.07	75.00	75.00	77.42	77.42	80.00
Recall	100	83.33	87.50	87.50	100	100	100.0
RMS error	0.439	0.636	0.546	0.499	0.460	0.448	0.356

Table 6 Results of various ML techniques using dataset D

Model	DT	ID3	CHAID	NBC	RI	RF	EN
Accuracy	77.42	80.65	83.87	83.87	83.87	77.42	86.67
Precision	84.62	84.00	84.62	86.96	84.62	79.31	85.51
Recall	91.67	87.50	91.67	83.33	91.67	95.83	100
RMS error	0.418	0.440	0.412	0.397	0.402	0.401	0.389

Experiment 2: In order to study if newly joined students' performance can be predicted accurately at the time of taking admission or not, ML models have been applied on dataset B and results have been shown in Table 4.

Experiment 3: In order to study the influence of social, demographic, and behavioral factors on the performance of students, ML models have been applied on dataset C and results have been shown in Table 5.

Experiment 4: In order to study the influence of academic factors only on students performance, ML models have been applied on dataset D and results have been shown in Table 6.

4 Results Discussion

The objective of the study to answer the research questions as discussed in the introduction.

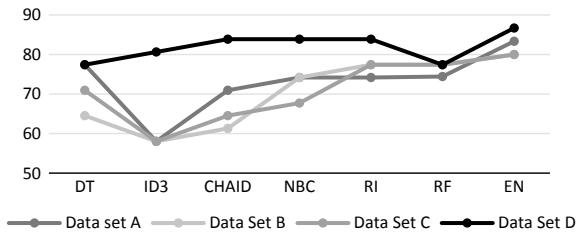


Fig. 2 Feature analysis using different datasets

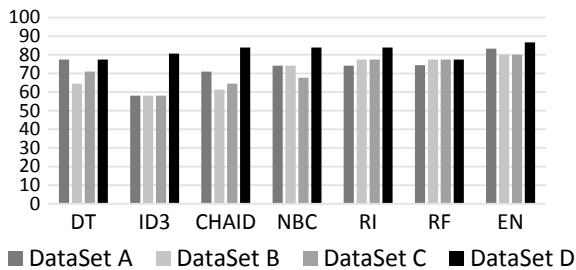


Fig. 3 Performance analyses of machine learning models

4.1 Influential Factors

For developing a generalizable model, it is important to identify the most influential factors that have predictive power. By comparative analysis, as shown in Fig. 2, it can be concluded that dataset D which contains all and only academic factors are the most dominating factors among others. Whereas dataset C which contains sociodemographic and economic factors also have a significant effect after academic factors on students' performance. It implies immediate corrective actions can be taken on the basis of academic attributes only.

4.2 Testing the Models Performance

The one purpose of this experiment is to analyze how effective these ML models are in identifying at-risk students. The Graph shown in Fig. 3 reveals that hybrid ML model Ensemble (Naive Bayes, Rule Induction, and random forest) perform best with all datasets followed by Rule induction and Random forest. Ensemble provides accuracy of 86.67 with 0.398 RMS error.

4.3 Early Prediction

It is recommended to predict at-risk students by including in between semester result marks. But early prediction can also be done with random forest or rule induction to provide instructors idea how students are going to perform. So low performer can get attention from starting. There is no need to wait for semester results. It provides a strong foundation of the course to the at-risk students which reduces students' early dropouts and improves their performance. To reduce the root mean square error and for improving the prediction accuracy, dataset A factors which contains academic factors along with social economic factors should be taken. This can improve surveillance periodically and then necessary corrective actions can be taken.

5 Conclusions

In this research, an effort is made to find the impact of different factors on students' performance prediction with the help of ML models. A feature space is constructed by considering academic, demographic, social, and behavioral factors. The features selection plays a vital role in selecting a subset of features which have more predictive power. The main objective was to improve the quality of higher education institutions by identifying students at risk so that concerned authorities may take preventive and corrective actions. It can be concluded from the results that apart from academic data, family qualification and income features have the significant impact on the performance of the student. Thus preventive actions may be taken by considering student performance, economic background, and family qualification. A hybrid approach of a model Ensemble, predicts students at risk most efficiently. On the other hand, immediate corrective actions can be taken on the basis of performance of students only to initially identify at-risk students to improve students' performance and reduce drop-out and failure rate.

Acknowledgments This research was supported by Chitkara University, Punjab, India. Authors would like to thank Chitkara University faculties and administration who provided academic data of students that greatly assisted the research. We also like to show our gratitude to the Chitkara University students who agreed and provided information to analysis demographic, social and behavior factors that greatly improve the result.

References

1. Scheuer, O., McLaren, B.M.: Educational data mining. In: Encyclopedia of the Sciences of Learning, pp. 1075–1079. Springer, Boston, MA (2012)

2. Farooq, M.S., Chaudhry, A.H., Shafiq, M., Berhanu, G.: Factors affecting students' quality of academic performance: a case of secondary school level. *J. Qual. Technol. Manag.* **7**(2), 1–14 (2011)
3. Thai-Nghe, N., Drumond, L., Krohn-Grimberghe, A., Schmidt-Thieme, L.: Recommender system for predicting student performance. *Procedia Comput. Sci.* **1**(2), 2811–2819 (2010)
4. Chatti, M.A., Dyckhoff, A.L., Schroeder, U., Thüs, H.: A reference model for learning analytics. *Int. J. Technol. Enhanc. Learn.* **4**(5–6), 318–331 (2012)
5. Baradwaj, B.K., Pal, S.: Mining educational data to analyze students' performance. (2012). [arXiv:1201.3417](https://arxiv.org/abs/1201.3417)
6. Lu, O.H., Huang, A.Y., Huang, J.C., Lin, A.J., Ogata, H., Yang, S.J.: Applying learning analytics for the early prediction of students' academic performance in blended learning. *J. Educ. Technol. Soc.* **21**(2), 220–232 (2018)
7. Aziz, A.A., Ismail, N.H., Ahmad, F.: MINING STUDENTS' ACADEMIC PERFORMANCE. *J. Theor. Appl. Inf. Technol.* **53**(3) (2013)
8. Baker, R.S.J.D.: Data mining for education. *Int. Encycl. Educ.* **7**(3), 112–118 (2010)
9. Sachin, R.B., Vijay, M.S.: A survey and future vision of data mining in educational field. In: 2012 Second International Conference on Advanced Computing & Communication Technologies (ACCT), pp. 96–100. IEEE (2012)
10. Zhang, Y., Oussena, S., Clark, T., Kim, H.: Using data mining to improve student retention in higher education: a case study. In: International Conference on Enterprise Information Systems (2010)
11. Borkar, S., Rajeswari, K.: Predicting students academic performance using education data mining. *Int. J. Comput. Sci. Mob. Comput.* **2**(7), 273–279 (2013)
12. Stephens, N.M., Fryberg, S.A., Markus, H.R., Johnson, C.S., Covarrubias, R.: Unseen disadvantage: how American universities' focus on independence undermines the academic performance of first-generation college students. *J. Pers. Soc. Psychol.* **102**(6), 1178 (2012)
13. Romero, C., Ventura, S., Pechenizkiy, M., Baker, R.S. (eds.): *Handbook of Educational Data Mining*. CRC press (2010)
14. Zaffar, M., Hashmani, M.A., Savita, K.S.: Performance analysis of feature selection algorithm for educational data mining. In: 2017 IEEE Conference on Big Data and Analytics (ICBDA), pp. 7–12. IEEE (2017)
15. Helal, S., Li, J., Liu, L., Ebrahimie, E., Dawson, S., Murray, D.J.: Identifying key factors of student academic performance by subgroup discovery. *Int. J. Data Sci. Anal.* 1–19 (2018)
16. Sivakumar, S., Selvaraj, R., Predictive Modeling of students performance through the enhanced decision tree. In: *Advances in Electronics, Communication and Computing*, pp. 21–36. Springer, Singapore (2018)
17. Adhatrao, K., Gaykar, A., Dhawan, A., Jha, R., Honrao, V.: Predicting students' performance using ID3 and C4. 5 classification algorithms (2013). [arXiv:1310.2071](https://arxiv.org/abs/1310.2071)
18. Ramaswami, M., Bhaskaran, R.: A CHAID based performance prediction model in educational data mining (2010). [arXiv:1002.1144](https://arxiv.org/abs/1002.1144)
19. García, E., Romero, C., Ventura, S., De Castro, C.: An architecture for making recommendations to courseware authors using association rule mining and collaborative filtering. *User Model. User-Adapt. Interact.* **19**(1–2), 99–132 (2009)
20. Breiman, L.: Random For. *Mach. Learn.* **45**(1), 5–32 (2001)
21. Pandey, M., Taruna, S.: A comparative study of ensemble methods for students' performance modeling. *Int. J. Comput. Appl.* **103**(8) (2014)
22. Pandey, M., Taruna, S.: An ensemble-based decision support system for the students' academic performance prediction. In: *ICT Based Innovations*, pp. 163–169. Springer, Singapore (2018)
23. Costa, E.B., Fonseca, B., Santana, M.A., de Araújo, F.F., Rego, J.: Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Comput. Hum. Behav.* **73**, 247–256 (2017)

Performance Analysis of FSO Link in Log-Normal Channel Using Different Modulation Schemes



Harmeet Singh and Amandeep Singh Sappal

Abstract Free Space Optics (FSO) is an evolving branch of communication in which light signal is transmitted through free space or atmosphere which may be turbulent in nature. Due to this turbulence, the signal deteriorates resulting in low SNR and high BER. This paper focuses on finding the relationships among BER, SNR, outage probability, and power margin with respect to E_b/N_0 (the ratio of energy per bit to the noise power spectral density) by considering Log-normal channel in FSO system. These relations are derived using different modulation techniques viz. BPSK, QPSK, and 16-QAM. It has been observed that the power margin required to enhance the power of distorted signal at receiver end is significant for E_b/N_0 less than 8 dB. Moreover, QPSK turns out to be the best modulation scheme among the three in the FSO system.

Keywords Log-normal · BER · SNR · Outage probability · Power margin · BPSK · QPSK · 16-QAM

H. Singh (✉) · A. S. Sappal
Punjabi University, Patiala, India
e-mail: harmeetsingh85.hs@gmail.com

A. S. Sappal
e-mail: sappal73as@yahoo.co.in

H. Singh
Chandigarh University, Mohali, India

1 Introduction

Free space Optics (FSO) is an upcoming and very important branch of the communication system which uses light as signal, travelling from transmitter to receiver through free space or atmosphere [1]. The medium of communication in FSO is usually atmosphere which may suffer from different conditions at different times. The system works well when the weather is clear. However, if it is not then the signal fades while travelling through it which results in loss or distortion of information. Therefore, the signal has to be modulated in some way or the other before transmission so as to reduce the loss [2–6]. There are a number of modulation schemes which can be implemented on the light signal before transmission. These include On-Off Keying (OOK), M-ary Quadrature Amplitude modulation (M-QAM), M-ary phase shift keying (M-PSK), and many more, depending on the preferences of bandwidth and power requirement [3].

Till date, the researchers have analysed the performance of FSO systems by considering various parameters which include Bit Error Rate (BER), Signal-to-Noise Ratio (SNR), Scintillation index, average received irradiance, outage probability, and power margin using different channel models and modulation techniques [7, 8]. However, a very important parameter viz. E_b/N_o (the ratio of energy per bit to noise power spectral density) has not been analysed in case of FSO which can give the idea of how much fading occurs due to noise when the signal is transmitted from transmitter to receiver through a turbulent channel. It is calculated at the input of the receiver to find out how strong the signal is. By considering this parameter, important information can be inferred from the analysis which can be used to attribute the amount of power required to be given to the signal for achieving a significant BER. The focus of this paper is to draw the significance of this parameter and analyse the system performance by considering Log-normal turbulent channel for transmission of FSO signal. The idea behind this work is to derive the relationship among BER, SNR, outage probability, and power margin for different values of E_b/N_o in FSO system to infer the amount of extra energy that needs to be supplied to the signal for achieving required BER at the receiver. Further, the outage probability for different modulation schemes (like BPSK, QPSK, and QAM) is derived in Log-normal turbulent channel.

The rest of the paper is systemized as follows: Sect. 1 contains a brief explanation of related work done by researchers, Sect. 2 derives the relationship between fading strength and E_b/N_o for different modulation schemes, Sect. 3 gives the importance of E_b/N_o in the relationship between BER and SNR in terms of signal strength, Sect. 4 shows the significance of E_b/N_o in the relationship between outage probability and power margin, Sect. 5 gives the performance analysis followed by conclusion in Sect. 6.

2 Relationship Between Fading Strength and E_b/N_0 for Different Modulation Schemes

In a communication system, the channel can be modelled as [9]

$$p = ix + j \quad (1)$$

where, x is the transmitted signal through channel after modulation, p is the received signal, i is the amplitude scaling factor and j is the Additive White Gaussian Noise having variance equal to σ^2 and mean equal to zero. AWGN is the basic noise model in which the signal is impaired with white noise which is linear in nature. In this channel, the noise variance and noise power spectral density (J_0) are related to each other as

$$\sigma^2 = \frac{J_0}{2} \quad (2)$$

M-ary modulation schemes, such as BPSK, QPSK, and QAM, have symbol energy given by

$$E_s = K_m K_c E_b \quad (3)$$

where, E_s = Symbol energy per modulated bit, $K_m = \log_2(M)$ with $M = 2$ for BPSK, $M = 4$ for QPSK, and so on, K_c is the code rate of the system which is usually equal to 1 and E_b is the energy per bit. Equation 3 can also be presented as

$$\frac{E_b}{N_0} = \frac{E_s}{K_m K_c N_0} = \frac{E_s}{K_m K_c 2\sigma^2} \quad (4)$$

where, E_b/N_0 is a ratio of energy per bit to noise power spectral density which is significant in computing the power of signal at receiver end. From Eq. 4, σ^2 comes out to be [9]

$$\sigma^2 = \left[2E_s K_m K_c \frac{E_b}{N_0} \right]^{-1} \quad (5)$$

2.1 For BPSK

In BPSK modulation, the binary input bits are generated in series in which 1 s are represented by +1 and 0 s by -1. The symbol energy per bit (E_s) for BPSK is equal to 1 and M is equal to 2. Therefore, K_m becomes $K_m = \log_2(M) = \log_2(2) = 1$. Considering these values, Eq. 5 for BPSK can be written as [9]

$$\sigma^2 = \left[2 \frac{E_b}{N_0} \right]^{-1} \quad (6)$$

2.2 For QPSK

In QPSK, a symbol contains a combination of 2 bits. As the name of modulation technique (i.e. Quaternary PSK) suggests, four phases of the signal are generated. The symbol ‘00’ is represented by 1 (0 degree phase rotation), ‘01’ by k (90 degree rotation), ‘10’ by -1 (180 degree) and ‘11’ by $-k$ (270 degree). The symbol energy per bit (E_s) for QPSK is equal to 1 and M is equal to 4. Therefore, K_m becomes $K_m = \log_2(M) = \log_2(4) = 2$. Considering these values, Eq. 5 for QPSK can be written as [9]

$$\sigma^2 = \left[4 \frac{E_b}{N_0} \right]^{-1} \quad (7)$$

2.3 For 16-QAM

M-QAM is a modulation technique in which information is embedded in amplitude as well as phase of the sinusoidal carrier. It is a combination of M-ASK, which corresponds to amplitude modulation, and M-PSK, which corresponds to phase modulation. In the case of 16-QAM, 16 signal points are equally distributed in four quadrants. The noise variance, in the case of 16-QAM, is given by [9]

$$\sigma^2 = \left[\frac{8}{\sqrt{10}} \cdot \frac{E_b}{N_o} \right]^{-1} \quad (8)$$

3 Importance of E_b/N_0 in Relationship Between BER and SNR in Terms of Signal Strength

When the signal transmits through the turbulent channel it experiences fading, resulting in a decrease in strength of signal received at the further end. This decrement can be analysed with the help of parameter known as E_b/N_0 . This parameter is mainly used in digital communication systems. The amount of fading occurs during transmission directly effects the BER which will determine the signal quality

at receiver end. By considering different values of this parameter, the relationship between BER and SNR can be deduced and the extra power required to achieve the significant BER level can be inferred for better system performance. The relation among BER, SNR, and E_b/N_0 are obtained in the following subsections for different modulation schemes.

3.1 Using BPSK Modulation

The coherent demodulator output in BPSK after passing through a low-pass filter is given by [10]

$$i_D(t) = \frac{RAIxg(t)}{2} + J(t) \quad (9)$$

where, I is irradiance, A is optical modulation index, J(t) is the additive noise, and R is photo-detector responsivity. The unconditional BER in Log-normal channel in BPSK can be given as [3]:

$$P_e \cong \frac{1}{\sqrt{\pi}} \sum_{i=1}^n w_i Q\left(\sqrt{X_0} \exp\left(X_1 \left[\sqrt{2}\sigma_1 x_i - \sigma_1^2/2\right]\right)\right) \quad (10)$$

where, w_i and x_i are the weight factors and zeros of an nth-order Hermite polynomial. Substituting σ^2 from Eq. 6 into Eq. 10, we get:

$$P_e \cong \frac{1}{\sqrt{\pi}} \sum_{i=1}^n w_i Q\left(\sqrt{X_0} \exp\left(X_1 \left[\sqrt{\left(\frac{E_b}{N_0}\right)^{-1}} x_i - \left(4\frac{E_b}{N_0}\right)^{-1}\right]\right)\right) \quad (11)$$

3.2 Using QPSK Modulation

As mentioned in [3, 10], the unconditional BER for M-PSK is given by

$$P_{ec} = \frac{2}{\log_2 M} Q\left(\sqrt{(\log_2 M)\gamma(I)} \sin(\pi/M)\right) \quad (12)$$

where, $\gamma(I)$ represents the electrical SNR. In the case of QPSK, $M = 4$. Therefore, Eq. 12 reduces to

$$P_{ec} = Q\left(\sqrt{\frac{1}{\sigma^2}} \sin(\pi/4)\right) \quad (13)$$

Substituting the value of noise variance from Eq. 7 into Eq. 13, we get

$$P_{ec} = Q\left(0.7071\sqrt{\left[4\frac{E_b}{N_0}\right]}\right) \quad (14)$$

3.3 Using 16-QAM Modulation

For M-QAM, the closed form solution for unconditional BER is derived in a similar manner as in case of M-PSK. Looking at the equation of unconditional BER for M-PSK and equation for conditional BER of M-QAM, the equation for unconditional BER of M-QAM is determined as [3, 10]

$$P_{ec} = \frac{2\left(1 - \frac{1}{\sqrt{M}}\right)}{\log_2 M} Q\left(\sqrt{\frac{3\log_2 M\gamma(I)}{2(M-1)}}\right) \quad (15)$$

After substituting $M = 16$, Eq. 15 can be rewritten as:

$$P_{ec} = \frac{3}{8} Q\left(4\sqrt{\frac{1}{10\sqrt{10}}\frac{E_b}{N_0}}\right) \quad (16)$$

4 Significance of E_b/N_0 in Relationship with Outage Probability and Power Margin

The parameter named outage probability is used to determine the probability of outaging of signal in case of large fading with average value of BER less than its threshold. It can be illustrated in relation with SNR as follows [11]:

$$P_{out} = P(P_m\gamma(I) < \gamma^*) \quad (17)$$

where $\gamma(I)$ is the electrical SNR per bit of a noise channel, γ^* is the threshold SNR with atmospheric turbulence equal to nil and m is the power margin, which depicts the extra energy provided to the signal in order to enhance its strength which has been depreciated by fading caused by atmospheric turbulence. P_{out} represents the probability of SNR becoming less than its threshold value, ultimately resulting in failure of the signal to reach at the receiver end. By using this expression, the extra power required to strengthen the signal can be calculated so as to achieve the desired SNR.

Not only the power margin, but energy per bit required by signal is also an important parameter to analyse the outage probability. As the power of the signal can be

enhanced by increasing its energy per bit, therefore, E_b/N_o plays a significant role in determining the power margin required at a particular value of P_{out} .

4.1 In Case of BPSK Modulation

Not only the power margin, but energy per bit required by the signal is also an important parameter to analyse the outage probability. As power of the signal can be enhanced by increasing its energy per bit, therefore, E_b/N_o plays a significant role in determining the power margin required at particular value of P_{out} . Equation 17 can be written in the form of E_b/N_o by using Eq. (6) as:

$$P_{out} = Q\left(\sqrt{2\frac{E_b}{N_o}} \ln P_m - \left(2\sqrt{2\frac{E_b}{N_o}}\right)^{-1}\right) \quad (18)$$

By putting the value of $Q(x) \leq 0.5 \exp(-x^2/2)$ (i.e. Chernoff upper bound), P_m or m (i.e. power margin) can be obtained from Eq. 18, as follows:

$$m = \exp\left(\sqrt{-2 \ln 2 P_{out} \left(2\frac{E_b}{N_o}\right)^{-1}} + \left(4\frac{E_b}{N_o}\right)^{-1}\right) \quad (19)$$

4.2 In Case of QPSK Modulation

In the case of QPSK, outage probability can be found out by deriving Eq. 17 using Eq. 7 and getting the result as follows:

$$P_{out} = Q\left(\sqrt{4\frac{E_b}{N_o}} \ln m - \left(2\sqrt{4\frac{E_b}{N_o}}\right)^{-1}\right) \quad (20)$$

Power margin in case of QPSK is found out to be:

$$m = \exp\left(\sqrt{-2 \ln 2 P_{out} \left(4\frac{E_b}{N_o}\right)^{-1}} + \left(8\frac{E_b}{N_o}\right)^{-1}\right) \quad (21)$$

4.3 In Case of 16-QAM Modulation

In the case of QAM, outage probability can be found out by deriving Eq. 12 using Eq. 8 and getting the result as follows:

$$P_{out} = Q\left(\sqrt{\frac{8}{\sqrt{10}} \frac{E_b}{N_o}} \ln m - \left(2\sqrt{\frac{8}{\sqrt{10}} \frac{E_b}{N_o}}\right)^{-1}\right) \quad (22)$$

Power margin in case of QPSK is found out to be:

$$m = \exp\left(\sqrt{-2 \ln 2 P_{out} \left(\frac{8}{\sqrt{10}} \frac{E_b}{N_o}\right)^{-1}} + \left(\frac{16}{\sqrt{10}} \frac{E_b}{N_o}\right)^{-1}\right) \quad (23)$$

5 Performance Analysis

The graphs of BER versus SNR and outage probability versus power margin are plotted for -4 to 20 dB values of E_b/N_o for BPSK, QPSK, and 16-QAM modulations. The Figs. 1, 2 and 3 are plotted between BER and SNR, and the number of inferences is drawn. At E_b/N_o value of 4 dB, the BER drastically decreases in QPSK, while a similar pattern is observed at almost 7 dB for BPSK and 6 dB for 16-QAM. However, for range of values from -4 to 2 dB, the spread of E_b/N_o curves in graphs increases from BPSK to 16-QAM to QPSK. This implies that BER falls more radically in QPSK at lower values of E_b/N_o followed by 16-QAM and then by BPSK. Furthermore, from

Fig. 1 The plot between BER and SNR for BPSK using a range of E_b/N_o

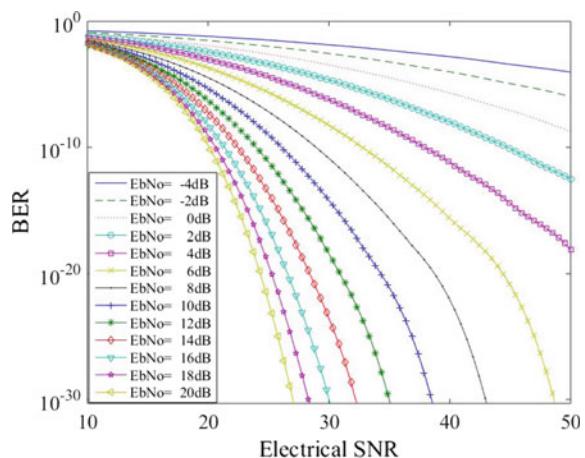


Fig. 2 The plot between BER and SNR for QPSK using a range of E_b/N_0

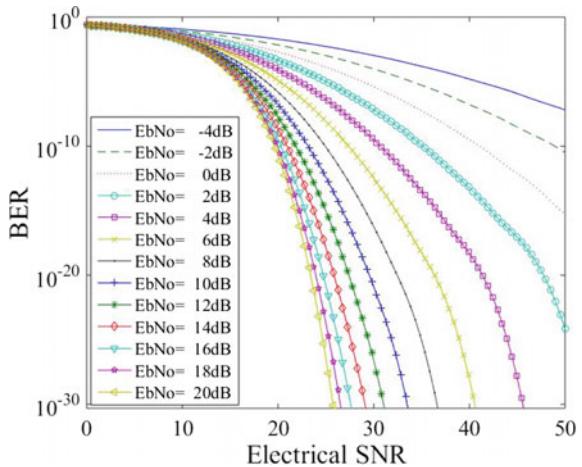
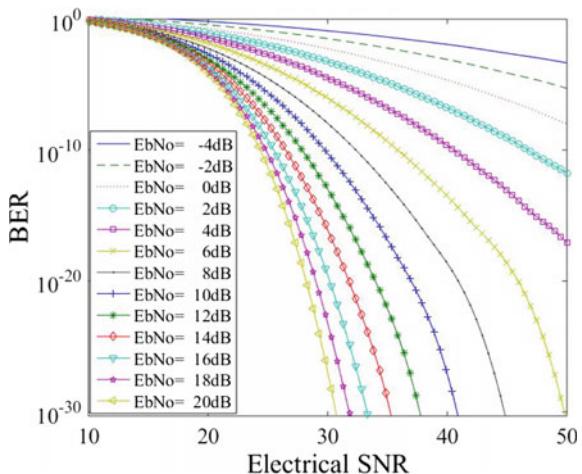


Fig. 3 The plot between BER and SNR for 16-QAM using a range of E_b/N_0



8 dB onwards the spread of the curves have a very small difference in all the three modulation techniques with QPSK having lowest BER among them.

In a nutshell, for a particular value of SNR, QPSK shows the least BER among the entire three modulation scheme, while BPSK has maximum BER followed by 16-QAM. Therefore, it can be inferred that QPSK is the best modulation scheme among the three. From the above observations, it can be deduced that if a signal with high energy is transmitted through a Log-normal channel, using all the three modulation schemes, then the QPSK modulation should be preferred over the other schemes.

Figures 4, 5 and 6, relate outage probability with power margin for range of E_b/N_0 values. Similar observations can be augmented by outage probability graphs for different values of E_b/N_0 . It can be analysed that for E_b/N_0 ranging from -4 to

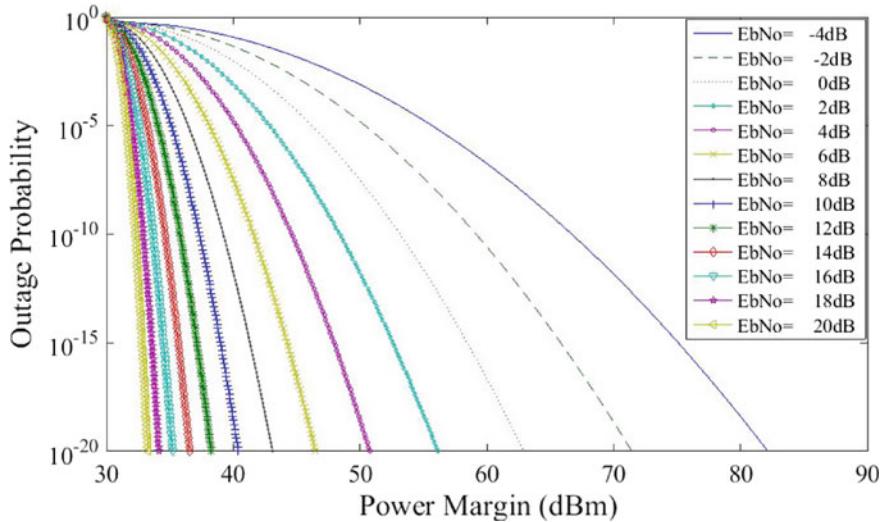


Fig. 4 The plot between Outage Probability and Power margin for BPSK using a range of E_b/N_0

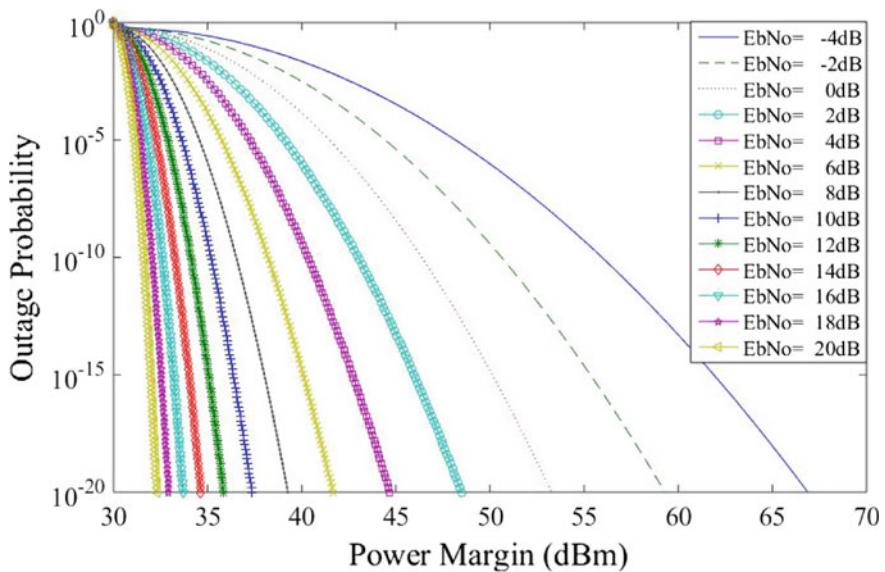


Fig. 5 The plot between Outage Probability and Power margin for QPSK using a range of E_b/N_0

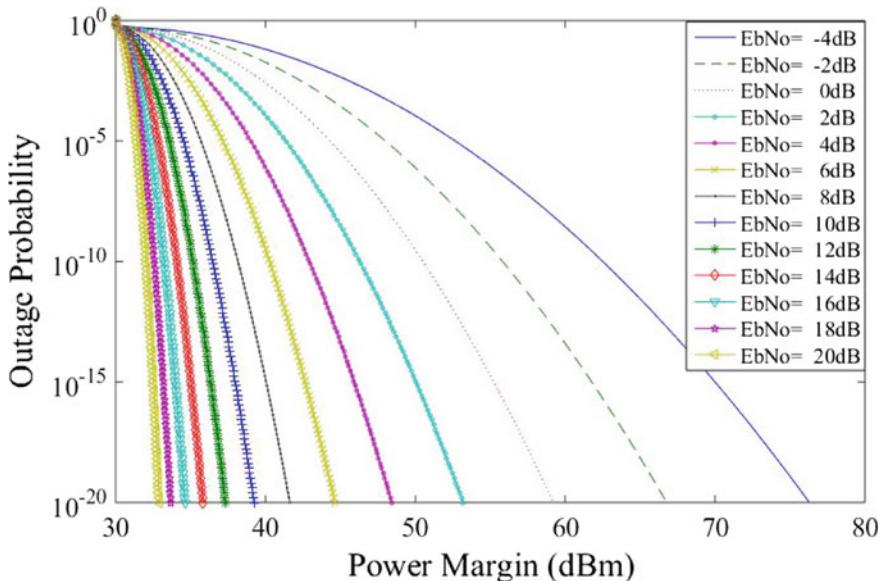


Fig. 6 The plot between Outage Probability and Power margin for 16-QAM using a range of E_b/N_0

2 dB, power required to prevent from out-ageing of signal is maximum in BPSK, pursued by 16-QAM and is least in QPSK. However, at 4 dB, the power margin values for all the three modulation schemes are almost the same with QPSK showing the least requirement of power.

It signifies that power margin or the extra power required to supply to the signal for achieving a sufficient signal strength at receiver end in order to avoid outage is found to be increasing in order of QPSK to 16-QAM to BPSK. This justifies the results inferred in points 1 to 4 of BER vs Electrical SNR of Log-normal channel. Therefore, in Log-normal channel, for transmission of optical signal in the FSO system, it should be preferably modulated using QPSK modulation technique than BPSK and 16-QAM. It is primarily due to the fact that noise variance in QPSK is comparatively very low than the other two as shown in Eq. (6) through Eq. (8). It can be supported by the fact that QAM modulation is the combination of amplitude and phase modulation, whereas, in M-PSK, only phase is altered. In FSO system, the

atmospheric turbulence directly affects the amplitude of the signal, thus degrading the performance of this modulation scheme.

6 Conclusion

In this paper, the significance of an important parameter E_b/N_o in determining the amount of signal power required to be supplied in order to achieve the necessary BER at receiver end is highlighted. The relationships among BER, SNR, outage probability, and power margin are determined with respect to E_b/N_o by considering Log-normal channel in the FSO system. The results are analysed for different modulation schemes viz. BPSK, QPSK, and 16-QAM. Various inferences are drawn from the analysis of graphs plotted between BER vs. SNR and Outage Probability vs. Power margin and it is concluded that QPSK performs better than the BPSK and 16-QAM in terms of BER and outage probability in Log-normal. The noise variance of 16-QAM is higher than BPSK and QPSK due to the fact that QAM modulation is the combination of amplitude and phase modulation. The atmospheric turbulence directly affects the amplitude of the signal, thus degrading the performance of this modulation scheme. For transmitting the signal with high energy through Log-normal channel using all the three modulation schemes, the QPSK modulation should be preferred over other schemes.

References

1. Willebrand, H., Ghuman, B.S.: *Free Space Optics: Enabling Optical Connectivity in Today's Network*. SAMS Publishing, Indianapolis (2002)
2. Ghassemlooy, Z., Hayes, A.R., Seed, N.L., Kaluarachchi, E.D.: Digital pulse interval modulation for optical communication. *IEEE Commun. Mag.* **36**(12), 95–99 (1998)
3. Ghassemlooy, Z., Popoola, W., Rajbhandari, S.: *Optical Wireless Communications: System and Channel Modelling with MATLAB*. CRC Press, Boca Raton (2013)
4. Mansour, A., Mesleh, R., Abaza, M.: New challenges in wireless and free space optical communications. *Opt. Lasers Eng.* **89**, 95–108 (2017)
5. Kaushal, H., Kaddoum, G.: Optical communication in space: challenges and mitigation techniques. *IEEE Commun. Surv. Tutor.* **19**(1), 57–96 (2017)
6. Al-Kinani, A., Wang, C.X., Zhou, L., Zhang, W.: Optical wireless communication channel measurements and models. *IEEE Commun. Surv. Tutor.* **20**(3), 1939–1962 (2018)
7. Li, J., Liu, J.Q., Tylor, D.P.: Optical communication using subcarrier PSK intensity modulation through atmospheric turbulence channels. *IEEE Trans. Commun.* **55**(8), 1598–1606 (2007)
8. Khalighi MA, Uysal M.: Survey on free space optical communication: a communication theory perspective. *IEEE Commun. Surv. Tutor.* **16**(4), 2231–2258 (2013)
9. Vishwanathan, M.: *Simulation of Digital Communication Systems Using MATLAB*. Published by Mathuranathan Vishwanathan at Amazon. (2013)

10. Proakis, J.G., Salehi, M.: Digital Communications, 5th edn. McGraw-Hill, New York (2007)
11. Chan, V.W.S.: Free space optical communications. *IEEE J. Light. Technol.* **24**(12), 4750–4762 (2006)

Performance Comparison of Uncoded and ZigZag Coded IDMA



Shiva Kaushik and Aasheesh Shukla

Abstract Wireless communication should be band limited and have the ability to provide low complexity at the receiver side, have to remain springy by reverence in the direction of information level, and offer modification compared to fading. So the scheme that provides all these things is interleave division multiple access (IDMA). IDMA can be considered as fifth generation nonorthogonal multiple access scheme (NOMA) and almost inherits all the advantages of code division multiple access (CDMA). In this paper the new zigzag coded based IDMA is presented and simultaneously the performance improvement is verified by comparing it by uncoded IDMA. Simulation experiments have been performed at different data length and Spreadlength establish the better BER performance of zigzag-coded IDMA.

Keywords IDMA · Zigzag code · BER · Decoding complexity · Encoder

1 Introduction

Zigzag codes are introduced by Ping et al. [1]. It is from the family of one-dimensional error correcting code technique which can be popularly used to reduce error in wireless communication [2]. This zigzag code is basically constructed from a zigzag graph. The geometrical structure of the code made the easy implementation of soft in soft out (SISO) decoding algorithm, which was less complex too, for the receiver. Zigzag codes also have low error floor which means it has less probability of flattening of curves at high SNR [3]. This code can be considered competitive with turbo code due to the complexity of decoding algorithm at receiver side which results in high error floor problem and it's also relatively weak at low coding rate [4]. Zigzag coding takes less time for simulation to analyze the performance of the system [5]. On the

S. Kaushik (✉) · A. Shukla

Electronics and Communication Engineering Department, G.L.A University, Mathura, U.P, India
e-mail: shiva.kaushik_mtec17@gla.ac.in

A. Shukla

e-mail: aasheesh.shukla@gla.ac.in

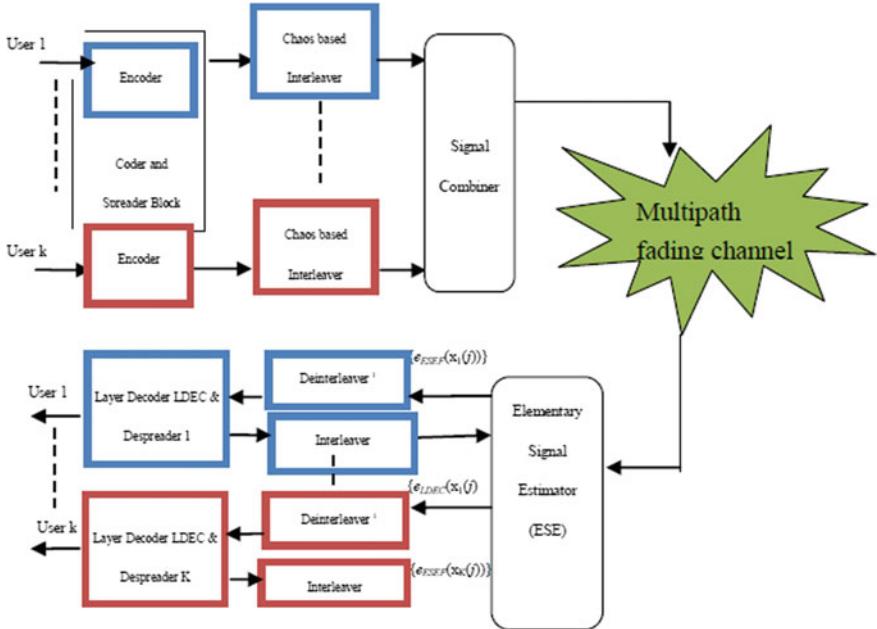


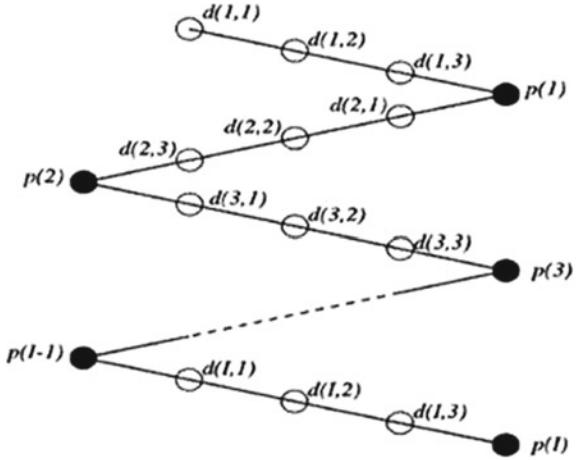
Fig. 1 Block diagram of transmitter and receiver structure of IDMA

other hand, a well-established multiple access is gaining popularity and is considered as a very robust NOMA scheme for 5G communication system termed as interleave division multiple access (IDMA), which also provide high spectral efficiency, less complex receiver, and better system performance. In this paper, zigzag code is used in combination with IDMA to improve the system performance. The paper is organized as follows, Sect. 2 provides the simple architecture of IDMA, and zigzag code is described in Sect. 3. The Sect. 4 presented the comparative analysis of zigzag-coded and uncoded and in the last Sect. 5 describes the conclusion and future scope of the work.

2 Interleave Division Multiple Access

Figure 1 shows the transmitter and receiver structure of interleave division multiple access (IDMA) system with K simultaneous users. The block diagram of IDMA system with K users is presented in Fig. 1. At the transmitter, the information bits $b_{i,k}$ are encoded by the encoder of rate R . The spreader code is simply a repetition code of rate R_r . The resulting code bits $c_{j,k}, j \in [1, \dots, N_b]$ are interleaved by user-specific interleaver Π_k to get $c'_{m,n,k}$, i.e., $c_k = [c_{k1}, c_{k2}, \dots, c_{kj}]$, where j is the chip length. Using the effective ISI channel the received signal can be represented as y_n .

Fig. 2 Highly structured zigzag graph



Here the transmitted symbol can be written as $\dot{x}_n \triangleq [x_{n1} \dots x_{nK}]^T$ and noise vector may be $\dot{\zeta}_n + \ell \triangleq [\zeta_n + \ell \dots \zeta_n + (\ell + 1)]^T$. In the receiver section, the received signal can be described as [13–15]:

$$R_j = \sum_{k=1}^K h_k x_{kj} + \zeta_j \quad (3)$$

3 Zigzag-Coded IDMA

A transmitter in IDMA adds parity bits to handled information from the source to protect the information. Basically the coding technique that we are going to implement with IDMA is known as zigzag coding. Zigzag code with four constituent encoders has low error floor than such a device that uses two constituent encoders. Zigzag belongs to the family of linear error correcting code used to reduce decoding complexity at the receiver side. Zigzag codes are defined by highly structured Zigzag graph. Due to its highly structured property of zigzag graph provides moderate complexity soft in/soft out (SISO) decrypt rule.

As shown in Fig. 2, the white and black nodes presented the input and parity bits such that input data bits: $d(j, k)st : j = 1, 2, 3 \dots J, k = 1, 2, 3 \dots K$. and parity of bits: $p(l), l = 1, 2 \dots L$.

4 Performance Analysis

To remove decoding complexity at the receiver's side in IDMA. Coding techniques are introduced to reduce decoding complexity at the receiver side generally known as coded IDMA. Coding technique that we are going to use is zigzag coding belongs to linear error correcting code introduce to reduce error and complexity at the receiver side of IDMA such type of IDMA are known as zigzag-coded IDMA. In zigzag-coded IDMA decode rule depends upon Max-Log-App (MAP) algorithm that requires only 20 additional equivalent operation (AEO) per information bit, per iteration. In this paper, a multiuser coded and uncoded IDMA communication device is examined. IDMA using various types of interleaver which can handle various types of data to improve performance with zigzag coding techniques. Zigzag coding techniques are also good in measuring complexity at the receiver side even at high complexity as well as high SNR in comparison to uncoded system providing a better result.

4.1 BER Analysis

To analyze the performance of any system simulation result plays a very important role to determine various characteristics of any system. As results show comparison between zigzag-coded IDMA versus uncoded IDMA which clearly states zigzag-coded IDMA provides better result in comparison to uncoded IDMA. Simulation result between BER vs. SNR (dB) clearly demonstrated the capacity of zigzag-coded IDMA to execute well in cellular environment for a multiuser communication. From simulation result we are going to analyze zigzag coding with IDMA by using random interleaver [6] to appraise BER of the system versus signal to noise ratio (SNR) in dB (Fig. 3).

5 Conclusion

From the simulation result, it can be easily concluded that zigzag-coded IDMA performs better in higher SNR. As we are going to a higher SNR, bit error rate (BER) decreases rapidly to improve channel capacity of IDMA. Simulation time of zigzag coding is less in comparison to another coding by using density-based evolution techniques to analyze the performance of IDMA system. The objective of this paper is to determine a better technique to reduce bit error rate (BER) at high SNR with the help of coding technique generally known as zigzag coding technique and compare it with uncoded system that decreases bit error rate (BER) at low SNR as shown in the above simulation result.

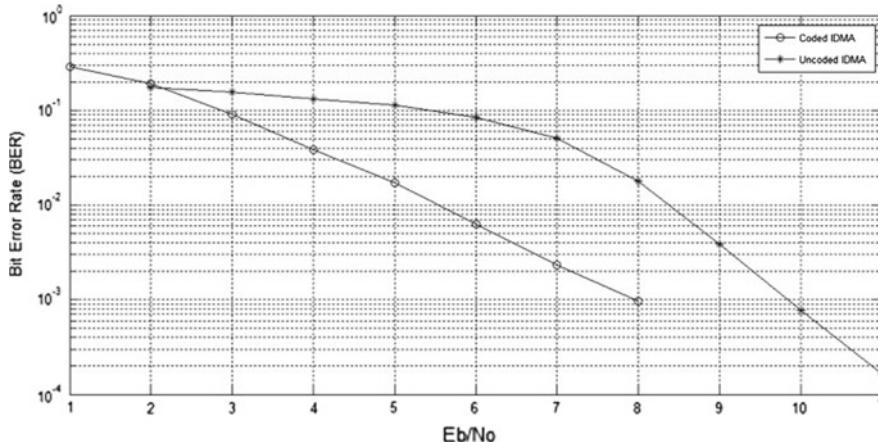


Fig. 3 Simulation results of zigzag-coded IDMA and uncoded IDMA

References

- Ping, L., Huang, X., Phamdo, N.: Zigzag codes and concatenated zigzag codes. *IEEE Trans. Inf. Theory* **47**(2), 800–807 (2001)
- Bilim, M., Kapucu, N., Develi, I.: Effect of repetition coding on the BER performance of interleave-division multiple access systems. *IEEE* **21**(4) (2015)
- Huang, X., Phamdo, N., Ping, L.: BER bounds on parallel concatenated single parity check arrays and zigzag codes. In: Global Telecommunications Conference, GLOBECOM '99, vol. 5, pp. 2436–2440 (1999)
- Zheng, W., tao Song, W., Zhang, H.-B., Liu, X.Z.: Design of concatenated turbo-SPC coded space-time systems. *J. Shanghai Univ. (English Edition)* **11**, 385 (2007). ISSN 1007-6417
- Berrou, C., Glavieux, A., Thitimajshima, P.: Near Shannon limit error correcting coding and decoding: turbo-codes. In: IEEE International Conference on Communications, 1993, ICC' 93 Geneva, Technical Program, Conference Record, vol. 2, pp. 1064–1070 (1993)
- Ping, Li, Liu, Lihai, Wu, K.Y., Leung, W.K.: Performance analysis of interleave division multiple access scheme with different coding techniques. *IEEE Trans. Wirel. Commun.* **5**(4), 938–947 (2006)
- Berrou, C., Glavieux, A.: Near optimum error correcting coding and decoding: turbocodes. *IEEE Trans. Commun.* **44**(10) (1996)
- Viterbi, Andrew J.: Very low rate convolutional codes for maximum theoretical performance of spread spectrum multiple-access channels. *IEEE J. Select. Areas Communication* **8**, 641–649 (1990)

Design of High Step-up Transformerless Modular Multilevel DC–DC Converter



Sonal Purkait, Rahul Pandey, Lalit kumar Sahu and Saji Chacko

Abstract The paper shows a structure that accomplishes high voltage proportions for interfacing sustainable power sources. The circuit topology accommodates high advance up DC–DC change proportions utilizing an MMC approach working in full mode. This topology works to venture up the input voltage with bigger change proportion. The MMC circuit framework comprises of an upper “N” and lower “M” set of cells. Phase-shift pulse width modulation (PS-PWM) is utilized to control voltage and power stream. PS-PWM with high obligation cycle is created to guarantee that every one of the capacitors is associated aside from one of them, which is out of the association. The recreations for the proposed topology are introduced. An open circle and shut circle arrangement framework is likewise spoken to for voltage guideline on the off chance that there is an adjustment in the input voltage. It is seen that using an anti-windup proportional integral controller, the output obtained is as desired and the saturation function helps to limit the voltage under certain specified values without compensating the voltage level. Thus, obtaining the conversion ratio obtained as 1:10.

Keywords We would MMC · SM · PS-PWM · FC · NPC · CHB

S. Purkait (✉) · R. Pandey · L. kumar Sahu · S. Chacko
Electrical Power System, SSTC, SSGI’s, Bhilai, India
e-mail: sonalpurkait1993@gmail.com

R. Pandey
e-mail: raahull17@gmail.com

L. kumar Sahu
e-mail: lkumar.ele@nitrr.ac.in

S. Chacko
e-mail: chackosaji68@gmail.com

1 Introduction

To structure a dc-to-dc converter has dependably been a test, i.e., the manner in which it ought to play out the transformation either unidirectional or bidirectional, conservativeness, control stream control, age of wide scope of yield levels, decrease in exchanging misfortunes, etc. [1, 2]. Among various DC–DC converters accessible, more degree for new topologies is in staggered converters. A one of a kind staggered converter that is having striking highlights like high versatility, particular nature, guideline of voltage, littler space necessities is Modular Multilevel Converters (MMC) [3, 4]. This paper sets a framework for examining Modular Multilevel Converters (MMC) to be connected and further produced for high proportion transformation.

A high voltage gain can be attained by cascaded boost, conventional boost, switched capacitor, and switched inductor converters or using multilevel topologies like neutral point clamped, flying capacitor and cascaded H-bridge. Limitations arise due to reverse recovery currents of diode resulting in reduced trustworthiness of the circuit. Also, while having high step-up voltage there might be electromagnetic interface concerns due to the reasonably high duty cycle levels [5]. Additionally, there is a restriction when utilized for high step-up voltage ratios because of the associated current spike may degrade component natural life or induce a failure [6]. These troubles can be solved by changing the turns ratio. Nevertheless, the said leakage of inductance energy in the transformer is caused by the switches having high voltage spikes. In accumulation, a high turns ratio increases the transformer losses, size, and cost. This results in reductions in system efficiency [7–11].

This paper displays a novel topology and control technique of a measured staggered transformerless DC–DC converter that can accomplish a high advance proportion dependent on the traditional lift with sub-modules put in both the diode and switch positions. Phase Shift PWM (PS-PWM) is utilized to achieve a high strong working frequency. The proposed converter has the component of measured quality, effortlessness, and adaptability.

1.1 High Step-up Transformerless DC–DC Modular Multilevel Converter

The configuration for high ratio step-up transformerless DC–DC MMC is shown below in Fig. 1. The circuit topology consists of two sections namely, upper sub-modules (N), and lower sub-modules (M). The number of upper SM is represented by “ $N = 2$ ” and the number of lower SM is given by “ $M = 4$ ”. In converting a simple standard switched mode circuit to modular multilevel format, a variety of capacitor clamped sub-modules are required.

The half bridge configuration in the upper SM has an IGBT with anti-parallel diode and a diode in series connection to provide unidirectional nature to the flow.

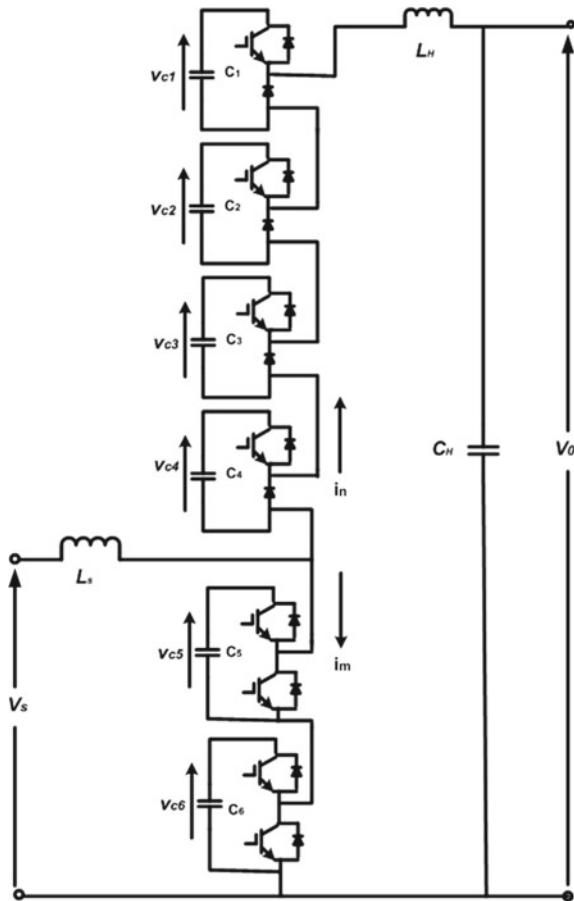


Fig. 1 A high step-up transformerless modular multilevel DC-DC converter

The lower SM has a half bridge configuration with two IGBTs with anti-parallel diode connected to each other. The lower SM IGBT's are used to charge the input inductor L_0 , and diode in the upper SMs are responsible to discharge current from high voltage side inductor L_H to high voltage end capacitor C_H . The output (high-side) voltage, V_0 is around equivalent to the aggregate of capacitor voltages of the pile of sub-modules once duty cycles are represented. There will be little contrasts between the quick voltage over the stack (as sub-modules switch) and the voltage over and across C_H .

1.2 Operation and Control

The analysis stresses upon the step-up conversion with a high ratio for which a clamping mechanism and phase-shifted pulse width modulation control is used. There are certain assumptions taken such as the switches are ideal and identical sub-modules are present, the converter isn't lossy, dc voltages are balanced, and the tank circuit parameters are greater than the input side parameters.

1.2.1 Phase-Shift PWM Control

Phase-shift pulse width modulation (PS-PWM) is used in this topology. The benefit of this modulation is that all the modules are equally used and low harmonics is produced even though the number of sub-modules is not high. A carrier set is owned by each arm. A high duty cycle is generated with PS-PWM to ensure one of the following conditions:

- At least three of the upper capacitors are connected in series with any one of the lower capacitors.
- Only the four upper capacitors are connected in series.

PS-PWM is orchestrated with a high obligation cycle to such an extent that just a single sub-module capacitor at once is out of the arrangement association and along these lines the progression up proportion of the circuit ends up reliant on the quantity of upper cells N. The successful frequency of this excitation is a lot higher than the frequency of exchanging of an individual cell. The proportional frequency of working is multiple times the benefit of exchanging frequency of the upper cells ($f = 4 fs$), and is two times the frequency of switching in the lower cells.

1.2.2 Operation Modes

In this topology there are eight modes, where the upper and lower voltages are complementary. The operating switching frequency T_e is indicated for these eight modes through waveform in Fig. 2 and is defined in Table 1.

Modes 1, 3, 5, 7—From Table 1, it is clear that modes 1, 3, 5, and 7 are the same. During these modes, all of the upper sub-module capacitors are connected to the resonant inductor, thereby creating a resonant tank as indicated in Fig. 3a. The output voltage is the addition of all the capacitor voltages in the upper sub-modules. Therefore, the output voltage for modes 1, 3, 5, and 7 can be written as:

$$V_0 = V_{C1} + V_{C2} + V_{C3} + V_{C4}, \quad (1)$$

where V_0 is the output voltage, and V_{C1} , V_{C2} , V_{C3} , and V_{C4} are the capacitor voltages for sub-modules 1 through 4, respectively.

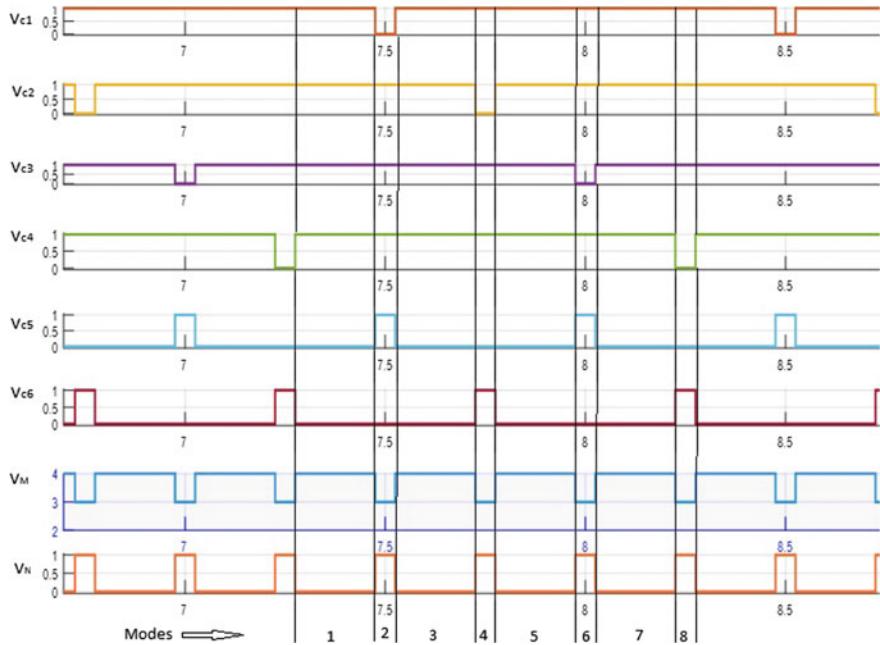


Fig. 2 Waveforms of the sub-module and output voltage of the high step-up converter

Table 1 Modes configuration during one period

	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5	Mode 6	Mode 7	Mode 8
Vc ₁	V ₀	0	V ₀					
Vc ₂	V ₀	V ₀	V ₀	0	V ₀	V ₀	V ₀	V ₀
Vc ₃	V ₀	0	V ₀	V ₀				
Vc ₄	V ₀	0						
Vc ₅	0	V ₀	0	0	0	V ₀	0	0
Vc ₆	0	0	0	V ₀	0	0	0	V ₀

Modes 2, 4, 6, 8—These modes can occur when one of the lower sub-module's capacitors is connected to the capacitors of the upper sub-modules as shown Fig. 3b. In other words, three of the upper capacitors are connected in series with one of the lower capacitors. The output voltages of the modes 2, 4, 6, and 8, respectively, can be written as:

$$V_{02} = V_{02} + V_{03} + V_{04} + V_{05}, \quad (2)$$

$$V_{04} = V_{01} + V_{03} + V_{04} + V_{06}, \quad (3)$$

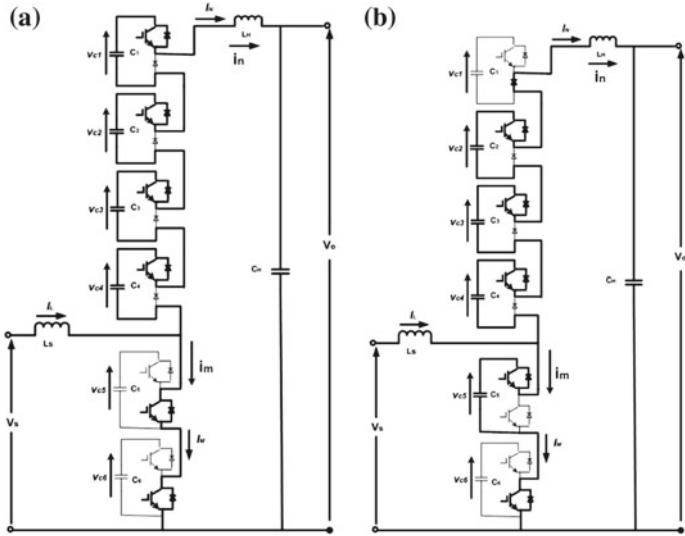


Fig. 3 Step-up operation modes in the first equivalent operating cycle **a** Mode 1. **b** Mode 2

$$V_{06} = V_{01} + V_{02} + V_{04} + V_{05}, \quad (4)$$

$$V_{08} = V_{01} + V_{02} + V_{03} + V_{06}, \quad (5)$$

where V_{02} , V_{04} , V_{06} , and V_{08} are the output voltages during modes 2, 4, 6, and 8, respectively. Based on (2), when V_{02} is absent V_{06} compensates for it. The same phenomena are happening for V_{05} , and V_{06} which compensate for the absent capacitor's voltage as indicated in (2) through (5). Consequently, by comparing (2) with (3), Eq. (6) results in:

$$V_{02} = V_{04} = V_{06} \quad (6)$$

Similarly, by comparing (3) with (5), Eq. (7) is written as:

$$V_{01} = V_{03} = V_{05} \quad (7)$$

1.3 Equivalent Circuit Analysis

As it was previously mentioned that modes 2, 4, 6, and 8 have the same on states. In addition, modes 1, 3, 5, and 7 have the same operational configuration. Therefore, operation of the transformerless MMC can be split into two modes. Select mode 1, that is the same as 3, 5 and 7, and mode 2, which is the same as 4, 6, and 8. To simplify

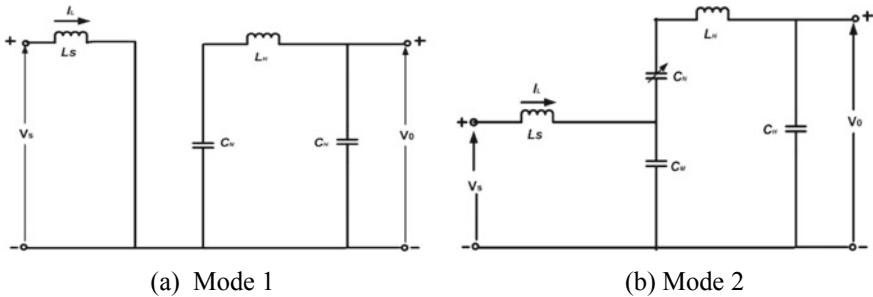


Fig. 4 Mode 1 & Mode 2 equivalent circuit of the transformerless MMC

In the theoretical analysis, suppose that the dc voltages in capacitor of all sub-modules are the same. When the IGBT in sub-module 1 is on, then mode 1 begins, and it ends with switching off the IGBT in sub-module 2. In mode 1, the input voltage is the same as the voltage of the low side inductor L . During this mode the circuit equivalent is as shown in Fig. 4a.

During mode 2, the capacitor in sub-modules 2 and 5 are out of the series connection, and the capacitor C_6 is connected with the capacitors of the upper sub-modules C_1 , C_3 , and C_4 , together with the resonant inductor L_r and the output capacitor C_H as shown in Fig. 4b. Accordingly, the resonant frequency can be written as:

$$f_r = \frac{1}{2\pi\sqrt{\frac{LrC}{N}}}$$

2 Simulation Results for Step-up MMC

The simulated results of the above designed DC-DC converter is shown with its response when into open-loop configuration and performance of the same parameters during the closed-loop configuration. In order to analyze the reports of the above two tests, the proper analysis of the modes, their occurrence, flow of current through the switches must be studied thoroughly. The conversion ratio is 1:10 for step-up, so as per given 30 V input the output is predictable to be 300 V. Cell voltage in the capacitors are clamped one by one in this mechanism. At any random time the number of capacitor voltages available is four. So, at the output summation of four capacitor voltages is found, it may be all the four from upper cells or three from the upper cell and any one from the lower power cell.

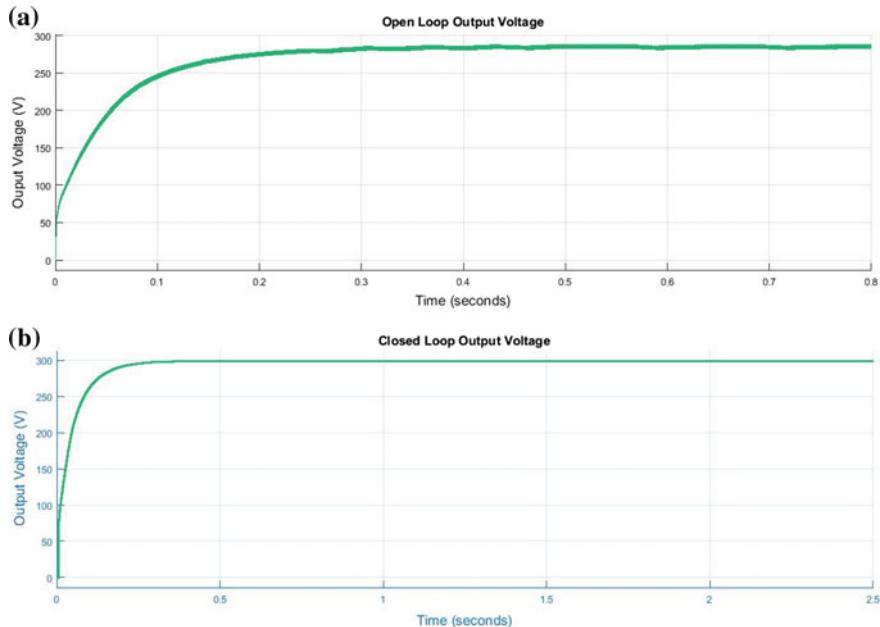


Fig. 5 MATLAB Simulation output voltage waveform of: **a** open-loop, **b** closed-loop system

2.1 Open-Loop Configuration

However with this configuration the output voltage obtained is smaller than ten times the value of voltage expected, i.e., the step-up ratio obtained is somewhat lesser than taken originally. There are conduction losses in the system that lead to this result, that is lowering the idealized value of voltage in output. The input voltage is selected to be 30 V and the output voltage is 282 V. The response of output is given away by Fig. 5a. In Mode 1, the low side voltage V_L is responsible to charge the current i_L of inductor L by the help of IGBTs in cells 5 and 6. When the upper sub-module current i_N in Mode 1 is crossing from zero to negative, damping of the current resonance occurs during the period the current is resonating freely. The reason behind it being the voltage drop caused by the switches. When i_N attains negative value in Mode 1, we obtain positive voltages across the diodes. Again when the transition of i_N occurs from negative to positive, voltage drops across the switches become negative. It is because the input side inductor is charging and the high voltage capacitor C_H is being discharged.

Until the end of this mode, the rise of upper sub-module current i_N is suppressed. When Mode 2 starts, discharging of the inductor current i_L in high voltage side occurs and current i_N continuously resonates until the end of the equivalent corresponding operating cycle is reached or the values drop to zero.

Table 2 Parameters of the simulation for closed-loop step-up MMC

Symbol	Quantity	Value
V_L	Low voltage dc input	30 V
V_H	High voltage dc output	300 V
C_H	High voltage region capacitor	180 μ F
C_L	Low voltage region capacitor	470 μ F
L_0	Low voltage side inductor	821 μ H
L_H	Output side inductor	120 μ H
R_H	High voltage side resistor	1070 Ω
R_L	Low voltage side resistor	18 Ω
T_e	Equivalent operating cycle	250 μ s
T_b	Sampling period	100 μ s
d	Duty ratio	0.64

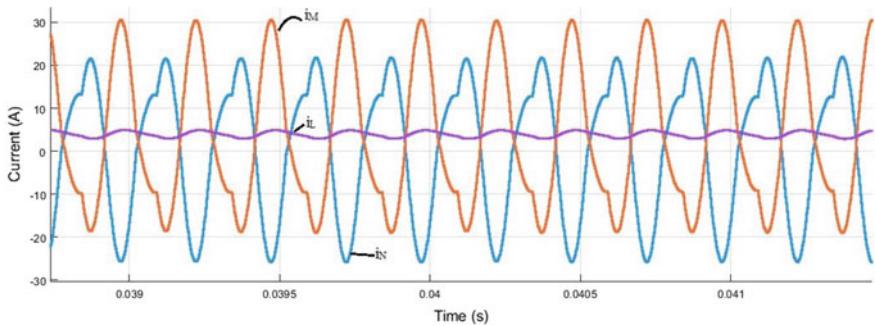


Fig. 6 Currents in the circuit: i_N (upper SM current), i_L (input current), i_M (lower SM current)

2.2 Closed-Loop Configuration

Inorder to test the regulation of the parameters of the converters, a closed-loop test is performed. The result in the form of waveform is obtained by MATLAB/SIMULINK.

The voltages obtained in the output is this time fed to a PI controller inorder to change the duty ratio. This is to imply that if the voltage on the high end side of the converter is worse than the estimated voltage value, it is the function of the PI controller to increase the effective duty cycle above the ideal value. The input low voltage is again 30 V as fixed. The closed-loop parametric values are listed below in Table 2. The duty ratio in closed-loop curve has been increased from 0.6 to 0.64 because of the controller. It can be evidently seen from the Fig. 5b, that the output waveform obtained is exactly as per the required value that is 300 V and around this value the high voltage is always controlled. The respective output currents waveform as obtained in closed-loop system is shown in Fig. 6.

Table 3 Comparison of parameters of open-loop test and closed-loop test

Parameters	Open-loop test	Closed-loop test
Low dc input voltage, Vs	30 V	30 V
High dc output voltage, Vo	282 V	300 V
Duty ratio, d	0.6	0.64

By comparing the voltages in the reference to each cell's DC voltage, for regulation a proportional feedback control unit is used. To limit the adjustable range of voltage in output a saturation function is used. The values of phase-lead control feedback parameters are $K_p = 1.92883615064021$ and $K_i = 84,176.9852090169$. If the voltages of capacitor are balanced for all the sub-modules and equal to the conversion ratio of voltage is to be:

$$\frac{V_0}{V_s} = \frac{N}{1 - d} \quad (8)$$

The voltage conversion ratio depends on both N and d. For example, if the duty cycle is 0.6 and the number of the upper cells is equal to 4, then the conversion ratio is approximately 1:10. By adjusting the duty cycle and choosing the suitable number of upper cells, the voltage conversion ratio can be 1:10.

2.3 Comparison of Test Results

In step-up MMC operation in open-loop, the stacking of voltages is obtained which is actually lesser than the estimated ideal assessment due to voltage drops in the switches. As, a result of which we infer to the conclusion that the step-up ratio is not 1:10 but somewhere lower (Table 3).

The duty ratio in closed-loop curve has been increased from 0.6 to 0.64 because of the controller. The output waveform obtained is exactly as per the required value that is 300 V. It is seen that the value of the voltage in output is higher in case of closed-loop system than in open-loop system.

3 Conclusion and Future Scope

This paper shows the effectiveness in using a new type of high ratio transformerless step-up DC–DC modular multilevel converter. This research shows that low voltage be it from renewable energy sources can be stepped up to almost ten times the value of input which is not only desirable, but achievable and realistic. The absence of the

transformer makes this circuit more cost-effective. Not only does the system have voltage conversion ratio in step-up as high, but it can be designed for bidirectional power flow by enabling step-down operation. Activities that could be explored for further developing the converter circuit are: fault analysis can be performed for MMC, state-space model for transient analysis and direct power control could be developed, bidirectional MMC design, etc.

References

1. Forouzesh, M., Siwakoti, Y.P., Gorji, S.A., Blaabjerg, F., Lehman, B.: Step-up DC–DC converters: a comprehensive review of voltage-boosting techniques, topologies, and applications. *IEEE Trans. on Power Electron.* **32**(12), 9143–9178 (2017)
2. Debnath, S., Qin, J., Bahrani, B., Saeedifard, M., Barbosa, P.: A review on operation, control, and applications of the modular multilevel converter. *IEEE Trans. Power Electron.* **30**(1), 37–53 (2015)
3. Perez, M.A., Bernet, S., Rodriguez, J., Kouro, S., Lizana, R.: Circuit topologies, modeling, control schemes and applications of Modular multilevel converters. *IEEE Trans. Power Electron.* **30**(1), 4–17 (2015)
4. Mhiesan, H., Al-Sarray, M., Mackey, K., McCann, R., Maher, M., Babaei, E., Sabahi, M., Hosseini, S.H.: High step-up/down ratio isolated modular multilevel for battery energy storage system on grid DC-DC converter. In: IEEE Green Technologies Conference (2016). ISSN: 2166-5478
5. Solas, E., Abad, G., Barrena, J., Arunetxea, S., Carcar, A., Zajac, L.: Modular multilevel converter with different submodule concepts-part I: capacitor voltage balancing method. *IEEE Trans. Ind. Electron.* **60**(10), 4525–4535 (2013)
6. Solas, E., Abad, G., Barrena, J., Arunetxea, S., Carcar, A., Zajac, L.: Modular multilevel converter with different submodule concepts-part II: experimental validation and comparison for HVDC application. *IEEE Trans. Ind. Electron.* **60**(10), 4536–4545 (2013)
7. Sahoo, M., Kumar, K.S.: High gain step up DC-DC converter for DC micro-grid application. In: 2014 7th International Conference on Information and Automation for Sustainability (ICIAFS), Colombo, pp. 1–5 (2014)
8. Liang, T.J., Lee, J.H.: Novel high-conversion-ratio high-efficiency isolated bidirectional DC-DC converter. *IEEE Trans. Ind. Electron.* **62**(7), 4492–4503 (2015)
9. Bhaskar, M.S., Sreeramula Reddy, N., Kumar, R.K.P., Gupta, Y.B.S.S.: A novel high step-up DC-DC multilevel buck-boost converter using voltage-lift switched-inductor cell. In: 2014 International Conference on Computer Communication and Systems, Chennai, pp. 271–275 (2014)
10. Choi, S., Agelidis, V.G., Yang, J., Coutellier, D., Marabeas, P.: Analysis, design and experimental results of a floating-output interleaved-input boost-derived DC-DC high gain transformerless converter. *IET Power Electron.* **4**(1), 168–180 (2011)
11. Hagiwara, M., Akagi, H.: Control and experiment of pulse width modulated modular multilevel converters. *IEEE Trans. Power Electron.* **24**(7), 1737–1746 (2009)

Mitigation of Haze Effects on Free Space Optical Communication Using Multibeam Technique



Gunjan Rathore, Charu Madhu and Anaahat Dhindsa

Abstract In a line of sight transmission of data, Free Space Optical communication (FSO) has served throughout history by carrying out communication through the atmosphere. Nowadays, the FSO is utilized as a speedy and simple solution for achieving maximum efficiency for broadband communication. The performance of FSO system has faced adverse consequences due to various changes in the properties of the channel such as weather conditions and length of the channel up to which the information needs to be transmitted. In this work, the performance of FSO system in Haze weather condition of Malaysia over a range of distance (i.e., from 30 to 1500 m) is studied. Better performance is achieved by the implementation of FSO communication in this weather by applying NRZ-Multibeam FSO and RZ-Multibeam FSO. Improved Q Factor is observed which demonstrates reduction in attenuation which is favorable to have a better communication through the system.

Keywords Bit error rate · Eye diagram · Free space optical communication · Simulation · Q-factor

1 Introduction

The requirement of high data rate for transmitting more data can be achieved through the enhancement of the optical communication mechanism. The transmitter, optical communication channel and receiver are the components of the Optical communication system. The electrical signal is converted into the light signal through the transmitter appliances. Further these light signals are carried to the receiver through the channel. There are several types of communication mechanisms like communication through diverse channels that are FSO [1], Optical Wireless Communication (OWC), and Optical Fiber. OWC is a technology that transmits information through an unguided channel which can either be free space or the atmosphere [2]. It is

G. Rathore (✉) · C. Madhu · A. Dhindsa

University Institute of Engineering and Technology, Panjab University, Chandigarh, India
e-mail: gunjanrathore93@gmail.com

further classified as indoor and outdoor OWC. The indoor OWC uses IR or visible light for communication within a building while the outdoor OWC also known as FSO uses laser beam for communication through the atmosphere. Currently several studies and analysis on the FSO communication are being carried out [3]. It is a sort of wireless RF communication technique. By applying the Radio over Fiber (RoF) technology like optical fiber, the FSO can offer high-speed point to point [4], point to multipoint communication that is otherwise difficult to achieve. The power consumption is directly linked with the distance and weather conditions resulting in a variation in efficiency [5]. It is executed by installing a laser device that can be placed in the buildings or offices and generally contain an optical transceiver between a laser transmitter and receiver to offer full duplex capability. For transferring the information a radio frequency spectrum, licensing is not required by the communication mechanism. It is thus utilized as a fast deployment communication mechanism. The expenses of execution are decreased by the FSO [6] and the full duplex communication is achieved [7].

The absorption, scattering, and beam divergence occurs in the FSO due to atmospheric channel [8] whose properties are a random function of space and time [9]. This makes it a random phenomenon which depends on the weather and geographical area. The particles and atoms spread in the atmosphere thus result in the atmospheric scattering due to the interaction of light with these particles. In propagation medium, the randomly distributed particles are generated under the pressure of atmospheric turbulence. The scattering and the multipath propagation of transmitted light is caused by the diverse refractive indices of these particles.

1.1 Weather Conditions

Different weather conditions like rain, fog, snow, storm, etc., severely affect the performance of FSO [10]. These weather conditions alter the composition of the atmosphere; thus, degrading the strength of the information signal in the form of light being carried. If smoke and suspended dust particles gather in moderately dry air then Haze is said to have occurred. In the atmospheric condition of Haze, the visibility is reduced and the degradation of the transmitted signal takes place in the optical communication system.

Fog. Fog is the major factor that causes attenuation when it comes to FSO. It consists of small droplets of water having radii near to the size of IR wavelengths. The weather conditions are said to be foggy when this brings a drop in the level of visibility that is the visibility lies between 0 and 2000 m. Attenuation here is expressed as

$$A = \frac{13}{V} \left(\frac{\lambda}{\lambda_0} \right)^{-q} \quad (1)$$

Here,

- A Atmospheric attenuation coefficient
- V Visibility (km)
- λ Wavelength (nm)
- λ_0 Reference wavelength = 550 nm
- q Particle size distribution

For analyzing the performance of FSO system the visibility data needs to be converted into the amount of attenuation. Firstly, installing a temporary system at the site for checking the visibility range and, then by using Kim and Kruse models to calculate the attenuation for the corresponding values of visibility. According to these models:

$$q = \begin{cases} 1.6 & V < 50 \text{ km} \\ 1.3 & 6 < V < 50 \text{ km} \\ 0.16V + 0.34 & 1 < V < 6 \text{ km} \\ V - 0.5 & 0.5 < V < 1 \text{ km} \\ 0 & V < 0.5 \text{ km} \end{cases} \quad (2)$$

Rain. In tropical areas, rain is a foremost factor affecting the visibility. For the rate of rain being R, the attenuation of optical wireless link is expressed as

$$A = kR^\alpha \quad (3)$$

Here k and α are power law parameters which depend on the amount of rain, rain temperature, and the size distribution of the raindrop. R is rain rate in mm/hr. The particular attenuation of the rain is expressed as:

$$A = \frac{2.9}{V} \quad (4)$$

Here, V = Visibility Range (km).

Air quality index and effects. To obtain the regular air quality information, Air Quality Index (AQI) is utilized. The AQI is centered on the problems related to health that can be experienced after inhaling the polluted air. For the air in any country to be clean the International Law regulates the foremost air pollutant particles suspended in the air by examining the AQI through the Environmental Protection Agency (EPA). At the ground level the particle pollution, nitrogen dioxide, carbon monoxide, and sulfur dioxide are involved in the ozone level. The attenuation due to Haze is expressed as follows:

$$\beta = \frac{3.91}{V} \left(\frac{\lambda}{550 \text{ nm}} \right)^{-q} \quad (5)$$

Here,

- β Haze attenuation,
- λ Wavelength (nm),
- V Visibility (km),
- q Distribution size of the particles.

2 Problem Formulation

Atmospheric conditions and climate changes affect each and every kind of communication process. The fiber optic communication highly suffers from the fluctuations in the weather. The climate such as snow, rain, smog, etc. affects the attenuation of the signals that travel through the free space optical channel. This issue needs to be resolved as the atmospheric fluctuations directly reduce the quality of data transmitted from source to destination. The objective of this work was to analyze the performance of multibeam FSO communication channel over a range of the visibility losses corresponding to the haze conditions in Malaysia. The drawback of the previous research was that it implements a single beam FSO for data communication which suffers from the issue of degradation in the performance of the system due to changes in the atmospheric conditions [11]. Thus there is a need to bring advancements in this. The consequences of the weather conditions can be analyzed through various parameters like Bit Error Rate (BER), Quality factor (Q) and Eye diagram [12].

3 Proposed Work

The previous research work had been done by implementing and analyzing FSO communication channel over a range of variations in the visibility losses in the haze [11]. To provide high data rate with high speed, a number of techniques have been proposed, but still no technique is ideally efficient. In this research paper, the single beam FSO is replaced with multibeam FSO channel. The reason behind this replacement is that the multibeam FSO communication channel has proved to be quite effective to reduce the temporary blockage that occurs due to obstruction whereas the single beam FSO fails to reduce this effect. The multibeam FSO is also found suitable at the transmitter side to facilitate the redundancy for the transmission path in the event where a particular laser source fails to do so. Along with the multibeam FSO channel, the proposed work also implements the analysis of an advance modulation scheme such as NRZ.

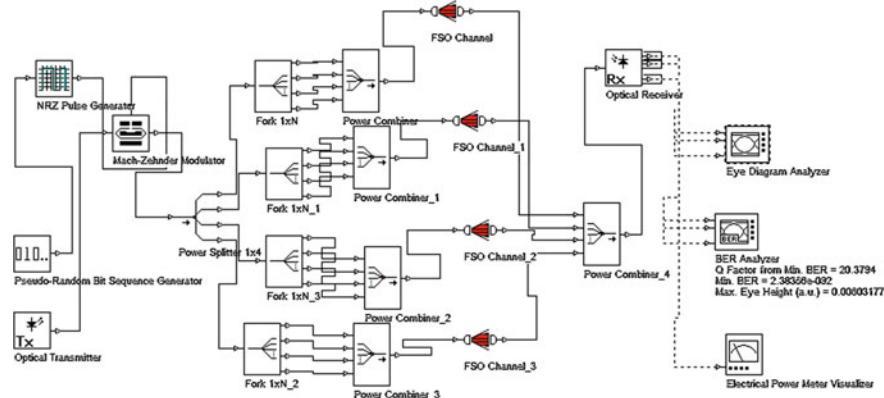


Fig. 1 Optisystem simulation setup in haze weather condition

4 Results

The testing, optimization, and designing of an optical link is accomplished in the Optical Communication Simulator (Optisystem). Consistent with the OSI model, the simulation is carried out from the physical layer to the transport layer. In FSO communication, the three essential components are the Transmitter (Tx), FSO Channel, and the Receiver (Rx). The Fig. 1 shows the setup in Optisystem. To generate the data that is modulated in the optical channel an external Pseudorandom bit sequence generator among Non-Return to Zero (NRZ) modulation is utilized. After this, the signal is transmitted once the FSO link length and the weather attenuation is set in accordance with the distance range (30–1500 m). The optical receiver (photodetector) was placed on the other side. The optical signal is converted into the digital signal by using the photodetector.

The values of BER, Q-factor and optical power are obtained from the BER analyzer, eye diagram analyzer, and the optical power meter. The effect of weather conditions on the signal being transmitted is analyzed by studying the output optical power at both transmitter and the receiver side by using a power meter. The quality of transmission was analyzed at the receiving end by the BER value and the eye diagram height. The Bit Errors were also examined throughout the transmission.

Comparison of the values of the proposed model with the NRZ modulation and RZ modulation to the base model is tabulated (Table 1).

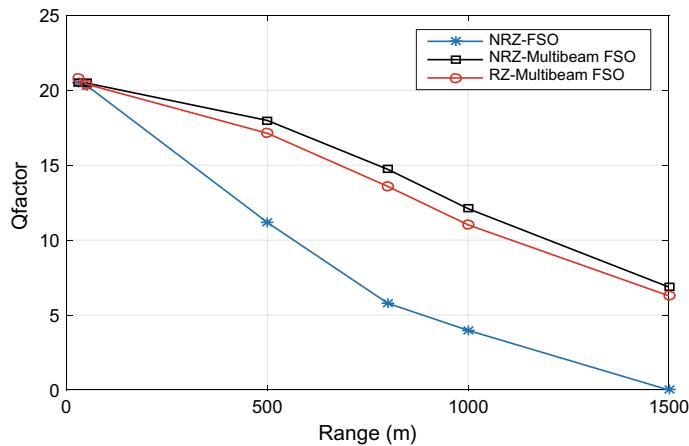
The graph of Fig. 2 depicts the Q Factor analysis of the proposed model. This graph shows that the results of the proposed model are better than the base model.

The comparison of the BER values of the proposed model with NRZ modulation and RZ modulation to the base model is tabulated (Table 2).

The Fig. 3 shows the BER analysis of the proposed model that attains better values as compared to the base model.

Table 1 Comparison of quality factor values

Weather condition	Visibility (in m)	dB/km loss	Base model	Proposed model with NRZ modulation	Proposed model with RZ modulation
Hazardous haze	30	93	20.508	20.543	20.819
Very unhealthy haze	50	55.75	20.319	20.483	20.405
Unhealthy haze	500	5.57	11.178	17.988	17.133
Sensitive haze	800	3.48	5.783	14.725	13.586
Moderate haze	1000	2.78	3.979	12.101	11.022
Good haze	1500	1.85	0	6.870	6.283

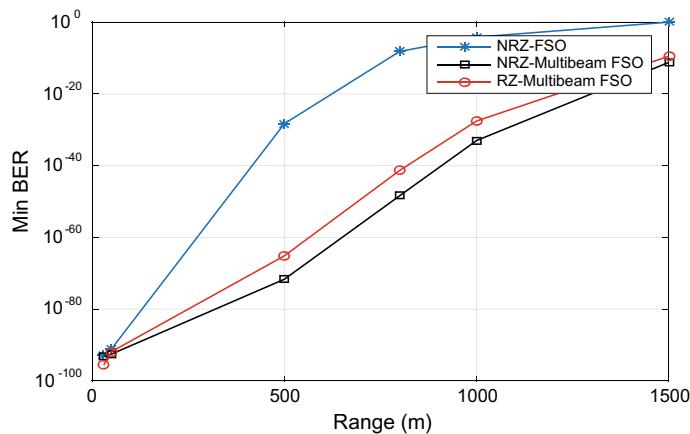
**Fig. 2** Q factor analysis

The comparison of the Eye Height values of the proposed model with NRZ Modulation and RZ modulation to the base model is tabulated (Table 3).

The graph of Fig. 4 depicts the analysis of the Eye Heights of the proposed model. In this graph, the values of the NRZ-Multibeam FSO (proposed model) and the values of RZ-Multibeam FSO (proposed model) are almost the same but better than the base model.

Table 2 Comparison of BER values

Weather condition	Visibility (in m)	dB/km loss	Base model	Proposed model with NRZ modulation	Proposed model with RZ modulation
Hazardous haze	30	93	1.693e-093	8.427e-094	2.594e-096
Very unhealthy haze	50	55.75	7.943e-092	3.035e-093	1.411e-092
Unhealthy haze	500	5.57	5.108e-029	2.239e-072	7.496e-066
Sensitive haze	800	3.48	7.024e-009	4.211e-005	4.614e-042
Moderate haze	1000	2.78	6.872e-005	1.007e-033	2.832e-028
Good haze	1500	1.85	1	6.146e-012	3.129e-010

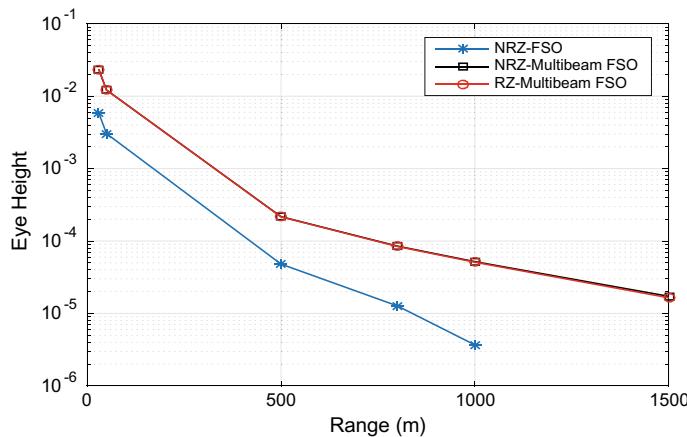
**Fig. 3** BER analysis

5 Conclusion

In the FSO-based WDM communication mechanism the environmental conditions have adverse effects. The WDM communication mechanism containing a single FSO channel was less useful and less proficient. Therefore a fresh method is offered here to control the attenuation through executing the multibeam WDM mechanism. The Optisystem simulations are carried out to analyze the performance of FSO by evaluating the parameters in Haze weather conditions of Malaysia at several distances. From the results, it is verified that the multibeam WDM communication system offered better results by applying NRZ and RZ modulation in comparison to the base

Table 3 Comparison of eye height values

Weather condition	Visibility (in m)	dB/km loss	Base model	Proposed model with NRZ modulation	Proposed model with RZ modulation
Hazardous haze	30	93	0.005	0.023	0.023
Very unhealthy haze	50	55.75	0.003	0.012	0.012
Unhealthy haze	500	5.57	4.803e-005	0.0002	0.0002
Sensitive haze	800	3.48	1.270e-005	8.500e-005	8.405e-005
Moderate haze	1000	2.78	3.670e-006	5.182e-005	5.093e-005
Good haze	1500	1.85	0	1.722e-005	1.644e-005

**Fig. 4** Eye height analysis

model which uses a single beam, in terms of the BER, Q Factor and Eye Height. The Quality Factor is increased (the minimum value being 6.28 as compared to 0 in the existing model) in this model by which the attenuation is decreased to achieve better communication. In Future, this research can be improved by replacing the proposed method with OCDM that is Optical Code Division Multiplexing technique and WDM.

References

1. patnaik, S., et al.: Free space optical communication: a review. *Int. Res. J. Eng. Technol.* **3**(4), 1261–1265 (2016)
2. Vaishali, Sancheti, S.: Practical analysis of bit error rate with respect to link distance in free space optical communication. *Int. J. Ind. Electron. Electr. Eng.* **5**(11) (2017)
3. Khalighi, M.A., Uysal, M.: Survey on free space optical communication: a communication theory perspective. *IEEE Commun. Surv. Tutor.* **16**(4), 2231–2258 (2014)
4. Malik, A., Singh, P.: Comparative analysis of point to point FSO system under clear and haze weather conditions. *Wirel. Pers. Commun.* 1–10 (2014)
5. Nor, N.A.M., Islam, M.R., Al-Khateeb, W.: Analysis of the effects of environment on the performance of free space earth-to-satellite optical link in Malaysia. In: IEEE Student Conference on Research and Development (SCOReD) pp. 235–240 (2011)
6. Rahman, T., Iqbal, S., Islam, M.M.: Modeling and performance analysis of free space optical communication system. In: International Conference on Informatics, Electronics & Vision (ICIEV), pp. 211–218 (2012)
7. Riza, N.A., Marraccini, P.J.: Power smart in-door optical wireless link applications. In: International Wireless Communications and Mobile Computing Conference (IWCMC), pp. 327–332 (2012)
8. Vitasek, J., Latal, J., Hejduk, S., Bocheza, J., Koudelka, P., Skapa, J.: Atmospheric turbulences in free space optics channel. In: 34th International Conference Telecommunications and Signal Processing (TSP), pp. 104–107 (2011)
9. Kaushal, H., Kaddoum, G.: Optical communication in space: challenges and mitigation techniques. *IEEE Commun. Surv. Tutor.* **19**, 57–96 (2017)
10. Fadhil, H.A., Amphawan, A., Shamsuddin, H.A.B., Hussein Abd, T., Al-Khafaji, H.M.R., Aljunid, S.A.: Optimization of free space optics parameters: an optimum solution for bad weather conditions. *Optik* **124**, 3969–3973 (2013)
11. Shah, S.M.A.: Effects of haze weather condition on 980 nm free space optical communication in malaysia. In: IEEE, pp. 1–7 (2016)
12. Garg, N., Singh, V.: Free space optical communication link using optical Mach-Zehnder modulator and analysis at different parameters. In: International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), 2014, pp. 192–195 (2014)
13. Khan, M.S., Muhammad, S.S., Awan, M.S., Kvicer, V., Grabner, m, Leitgeb, E.: Further results on fog modelling for terrestrial free-space optical link. *Opt. Eng.* **51**(3), 1–9 (2012)
14. Kvicer, V., et al.: Validating relationship between aerosol's liquid water content and optical attenuation for terrestrial FSO links. In: Telecommunications (ConTEL). In: Proceedings of the 11th International Conference, pp. 195–198 (2011)
15. Wu, Z., Chau, J., Little, T.: Modeling and designing of a new indoor free space visible light communication system. In: 16th European Conference Networks and Optical Communications (NOC), pp. 72–75 (2011)
16. Yu, B.S., Yao, Y., Zhao, Y.F., Liu, C., Yu, X.L.: Simulation research of medium-short distance free-space optical communication with optical amplification based on polarization shift keying modulation. *Optik* **125**, 3319–3323 (2014)

Spectrum Sensing in MIMO Cognitive Radio Networks Using Likelihood Ratio Tests with Unknown CSI



Juhি Singh and Aasheesh Shukla

Abstract Spectrum scarcity increases day by day due to underutilization of spectrum, Cognitive Radio Network (CRN) has been considered as a promising solution for in 5G communication system, as all of the available frequencies are not properly occupied in the spectrum. This is also very common to see that some of the frequencies in band are not utilized and some are overutilized. So the spectrum sensing is necessarily required to improve the accuracy of spectrum utilization and to protect the transmission of primary users. However, the dynamic nature of spectrum makes the sensing as a cumbersome task. In this paper, the problem of spectrum sensing in MIMO cognitive radio networks (CRN) has been considered in the existence of channel state information (CSI) uncertainty and various popular likelihood ratio test (LRT's) has been suggested and analyzed for spectrum sensing. The expressions for probability of false alarm and probability of detection are also obtained in closed form. Simulation experiments are performed in MATLAB to compare the performance of suggested LRT's and to identify the optimum LRT for CRN.

Keywords Cognitive radio · Spectrum sensing · Likelihood ratio tests (LRTs) · Energy detection · 5G communication

1 Introduction

Recently, it is observed that, to meet the requirements of advanced wireless applications of 5G communication, the demand for higher data rates and high-speed communication has been significantly increased. Simultaneously, the problem of scarcity of frequency spectrum is also increased due to the wide growth in advanced applications and congestion of network users. Also, there is wastage of frequency bands

J. Singh (✉) · A. Shukla

Department of Electronics and Communication, GLA University, Mathura, India
e-mail: juhisi.94@gmail.com

A. Shukla
e-mail: aasheesh.shukla@gla.ac.in

due to pre-allocation [1–3]. On the contrary call, it is interesting to note that, in many available research reports, and this is revealed that the paucity of spectrum arises, not because of unavailability of spectrum but due to inappropriate usage of existing spectrum bands. More specifically, in static band allocation the main policy is that the primary users (PU) have fixed dedicated channels and any other or secondary users (SU) have no permission to interfere in the bands even if the channel is vacant or unutilized. However, without causing any interference to the primary user recently proposed cognitive technology allowed the secondary users to get the opportunity to access the licensed spectrum [4, 5] or to get spectrum holes for using vacant bands. In the literature, many spectrum sensing schemes have been proposed. The basic definition of cognitive radio is appropriate spectrum sharing or reusing by allowing the secondary users to communicate in the frequency band allotted to primary user, only when the band is free or unused. So the secondary user (SU) continuously need to sense the spectrum of PU to find the vacant band for information transfer without creating interference to PU [6]. On an interesting note, the sensing becomes more crucial because when PU becomes active then SU has to vacate the band or decreased its transmit power. There are many challenges associated with cognitive radio such as; the signal to noise ratio (SNR) may be low at the detector. If the SU unit is few meters away, then SNR can be as low as -20 dB [3]. Multipath fading, time dispersion uncertainty, and noise power fluctuation make detection unreliable. Due to these issues, spectrum sensing becomes a more important research area for modern and future wireless communication. The popular sensing methods, which are proposed in the literature are; likelihood ratio test (LRT), energy detection (ED). Matched filter detection (MFD) and some other as Eigenvalue detection. Wavelet-based sensing, etc. First three, i.e., LRT, MFD, and ED require signal power as well as noise power for detection. Further, in this paper, the MIMO wireless system is used to increase the performance of spectrum sensing.

The main objective of this paper is to review the popular optimal one, without compromising with other resources. The rest of the paper is organized as follows. The system model for spectrum sensing is briefly described in Sect. 2. The brief reviews of spectrum sensing algorithms are discussed in Sect. 3. The simulation experiment analysis is explained in Sect. 3, and in Sect. 4 the conclusion of the work is presented. In the system model of MIMO it is assumed that a number of antennas at the receiver is $M \geq 1$. Now for the appropriate detection, there may be two hypotheses. For example, distinguishing between $-1, +1$ based on the observation at the receiver in BPSK modulation-based communication system. Each of these problems is a hypothesis testing problem, i.e., deciding or making an intelligent decision on which hypothesis is true. Specifically, the above scenario in which one is deciding between two hypotheses is termed as a binary hypothesis testing problem. Consider the problem given below with two hypotheses H_0 , and H_1

$$H_0 : \mathbf{x} = \mathbf{A}_0 + \mathbf{W}$$

$$H_1 : \mathbf{x} = \mathbf{A}_1 + \mathbf{W}$$

where \mathbf{W} is the Gaussian noise with $\mathbf{W} \sim \mathcal{N}(0, \sigma_w^2)$, i.e., hypothesis H_0 corresponds to Gaussian with mean A_0 while hypothesis H_1 corresponds to Gaussian with mean A_1 . The hypothesis H_0 is termed as “*null*” hypothesis and hypothesis H_1 is termed as “*alternative*” hypothesis. Generally the value of A_0 can be assumed as 0, i.e.,

$$\begin{aligned} H_0 : \mathbf{x} &= \mathbf{W} \\ H_1 : \mathbf{x} &= \mathbf{A} + \mathbf{W} \end{aligned}$$

The prime objective of spectrum sensing is to decide the appropriate hypothesis (choose H_0 or H_1) depending upon the received signal. The performance parameters for spectrum sensing are P_d and P_{fa} . P_d is known as probability of detection and mainly used to define the hypotheses H_1 (i.e., the algorithm correctly detecting the presence of primary user(PU) in the band) and P_{fa} is described as probability of false alarm which defines the hypotheses H_0 (i.e., the prediction for the presence of primary user (PU) in the band is mistakenly declared). The sensing methods or algorithm can be said “optimal” if P_d achieves the highest value for given P_{fa} .

1.1 False Alarm and Miss Detection

In spectrum sensing, false alarm can be defined as *type 1* error and miss detection can be understood as *type 2* error. These both parameter can be summarized as p_{fa} : Probability of false alarm that is probability of deciding H_1 when H_0 is true p_d : Probability of miss Detection that is probability of deciding H_1 when H_1 is true. There must be a proper tradeoff between these two probabilities so the probability of detection should be characterized as a function of the probability of false alarm. If γ is defined as a threshold value for the purpose of detection then P_{fa} and P_d can be written as

$$p_d = Q\left(\frac{\gamma - A_1}{\sigma_w}\right) \quad (1)$$

$$p_{fa} = Q\left(\frac{\gamma - A_0}{\sigma_w}\right) \quad (2)$$

$$\gamma = \sigma Q^{-1}(p_{fa}) + A_0 \quad (3)$$

where A_0 and A_1 are the respective mean value.

Therefore, from Eqs. (1) and (2) p_d in terms of p_{fa} is given as

$$p_d = Q\left(Q^{-1}(p_{fa}) + \frac{A_0 - A_1}{\sigma_w}\right)$$

1.2 Receiver Operating Characteristic

p_d versus p_{fa} relation is termed as receiver operating characteristic(ROC). Generally it is desirable to maximize the p_d for a particular value of p_{fa} .

1.3 Minimum Probability of Error

Probability of error is an important qualitative parameter and the sensing algorithm should have capacity to minimize the probability of error. In binary hypothesis testing error occurs if detected hypotheses is H_1 for H_0 or decide H_0 for H_1 . Therefore, the net probability of error is,

$$P(\text{dec } H_1 \cap H_0) + P_r(\text{dec } H_0 \cap H_1)$$

Observe that the minimum probability of error depends on the prior probabilities $p(H_0)$ and $p(H_1)$. The minimum probability of error framework is therefore Bayesian in nature. Let R_1 be the region corresponding to deciding H_1 , i.e., decide H_1 if x lies R_1 . So, the probability of error is,

$$p_e = p(H_0) * \left(\int_{R_1} p(x/H_0) dx \right) + p(H_1) * \left(\int_{R_1} p(x/H_1) dx \right)$$

p_e is minimized if R_1 is such that

$$\begin{aligned} (p(H_0) * p(x/H_0)) &\leq (p(H_1) * p(x/H_1)) \\ \Rightarrow H_1 \text{ if } \left(\frac{p(x/H_1)}{p(x/H_0)} \right) &\geq \left(\frac{p(H_0)}{p(H_1)} \right) = \gamma \\ \text{i.e., LRT With } \gamma &= \frac{p(H_0)}{p(H_1)} \\ \Rightarrow H_1 \text{ if } \left(\frac{p(H_1) * p(x/H_1)}{p(H_0) * p(x/H_0)} \right) &\geq 1 \end{aligned}$$

Observe

$$\begin{aligned} p(H_0/x) &= \left(\frac{(p(x/H_0) * p(H_0))}{(p(x/H_0) * p(H_0)) + (p(x/H_1) * p(H_1))} \right) \\ p(H_1/x) &= \left(\frac{(p(x/H_1) * p(H_1))}{(p(x/H_1) * p(H_1)) + (p(x/H_0) * p(H_0))} \right) \end{aligned}$$

Therefore,

$$\frac{p(H_1/x)}{p(H_0/x)} = \frac{p(x/H_1)P(H_1)}{p(x/H_0)p(H_0)}$$

2 Likelihood Ratio Test

One can therefore make one intelligent choice by looking at the likelihood of x corresponding to each hypothesis. The likelihood corresponding to observation x under hypothesis H_0 is

$$p(x; H_0) = \frac{1}{\sqrt{2\pi\sigma_w^2}} e^{-\frac{1}{2\sigma_w^2}(x-A_0)^2}$$

Similarly, the likelihood under hypothesis H_1 is

$$p(x; H_1) = \frac{1}{\sqrt{2\pi\sigma_w^2}} e^{-\frac{1}{2\sigma_w^2}(x-A_1)^2}$$

Therefore one can choose the hypothesis, corresponding to the greater among the two likelihoods, i.e.,

$$\begin{aligned} H_0 : p(x; H_0) &\geq p(x; H_1) \\ H_1 : p(x; H_0) &< p(x; H_1) \end{aligned}$$

Both the conditions can be summarized as a single condition by

$$H_0 \text{ if } \frac{p(x; H_0)}{p(x; H_1)} \geq 1$$

i.e., choose H_0 if the likelihood ratio

$$\frac{p(x; H_0)}{p(x; H_1)} \geq 1$$

Otherwise choose H_1 . This is termed as likelihood ratio test (LRT).

3 Results and Discussion

The main objective of this paper is to observe the performance of spectrum sensing through various methods like maximum likelihood ratio, Neyman–Pearson method, Match filter detection method, and lastly, Energy detection method by taking proper threshold value. For the number of Monte Carlo iterations $K = 10^5$, the author

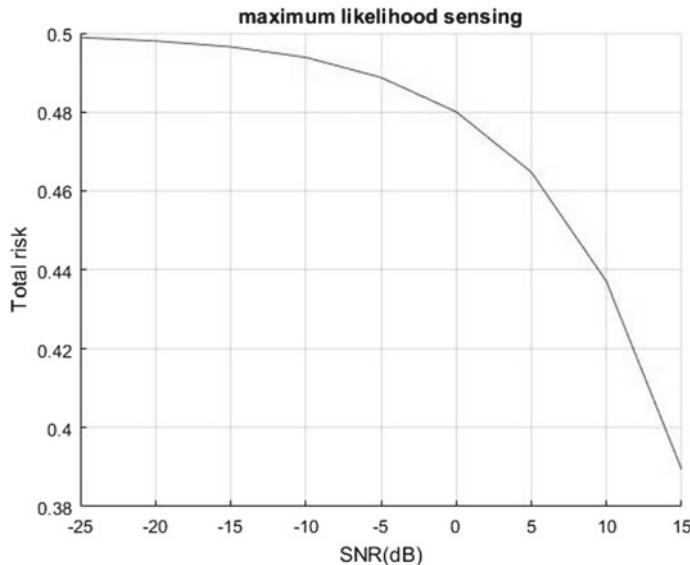


Fig. 1 ROC between probability of error or total risk and SNR of maximum likelihood sensing for SNR from -25 to 5 dB for $K = 10^5$ iterations

compares probability of detection versus probability of false alarm for different SNR and by taking equal value of threshold for all methods for above-mentioned methods. Through results it is shown that as on increasing the values of SNR the ROC curve reaches toward its ideal value means performance of detection increases (Figs. 1, 2, 3 and 4).

4 Conclusion

From above simulation results, it is clear that the performance of detector increases as SNR increases. Hence ROC is utmost important for analysis of detector for both the systems like Coherent and Non-Coherent systems. However, it is important to know the information about primary user's signal and its arrival time. Through spectrum sensing in CR, there is scope of proper utilization of spectrum for various future applications like IoT and 5G network NOMA.

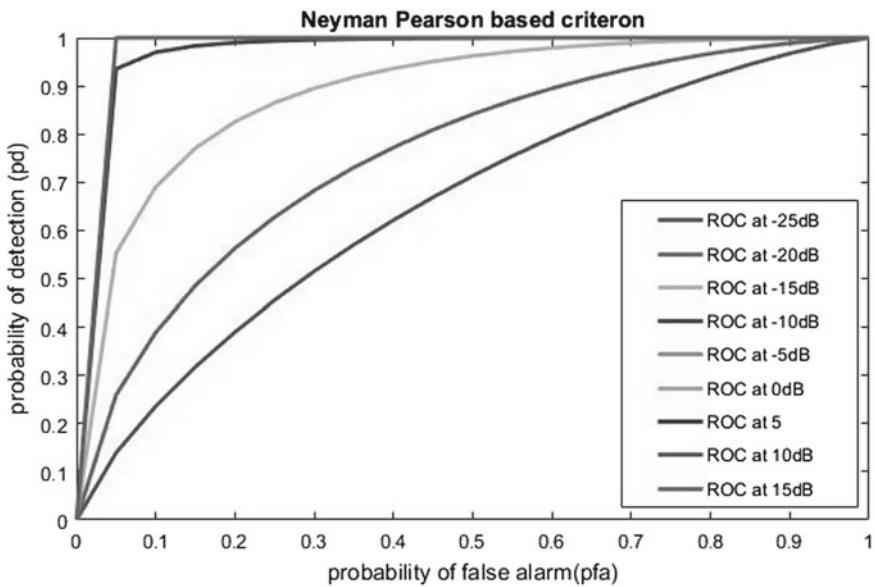


Fig. 2 ROC between probability of detection and probability false alarm of Neyman–Pearson criterion based sensing for SNR from -25 to 5 dB for $K = 10^5$ iterations

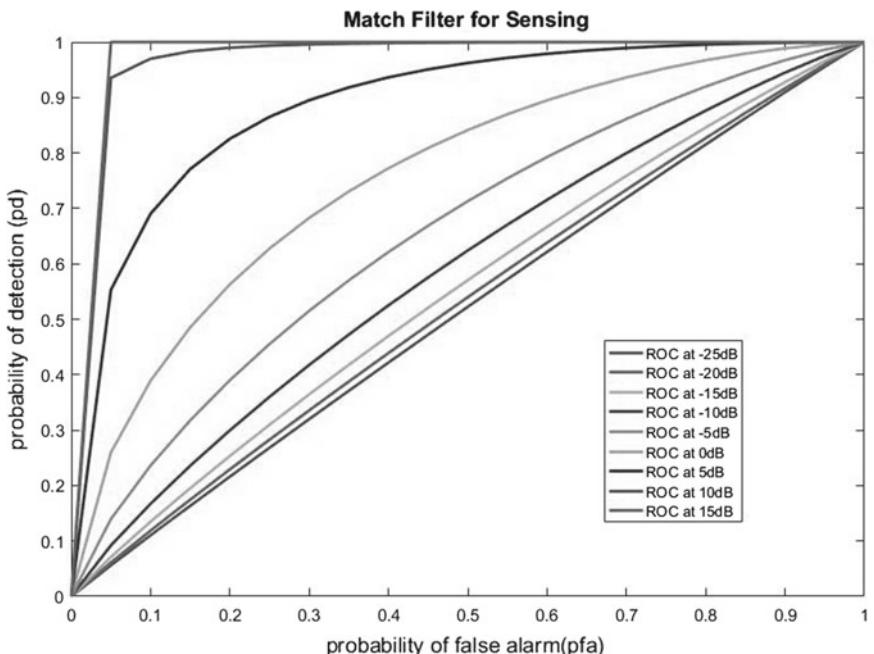


Fig. 3 ROC between probability of detection and probability of false alarm of Match filter sensing for SNR from -25 to 5 dB for $K = 10^5$ iterations

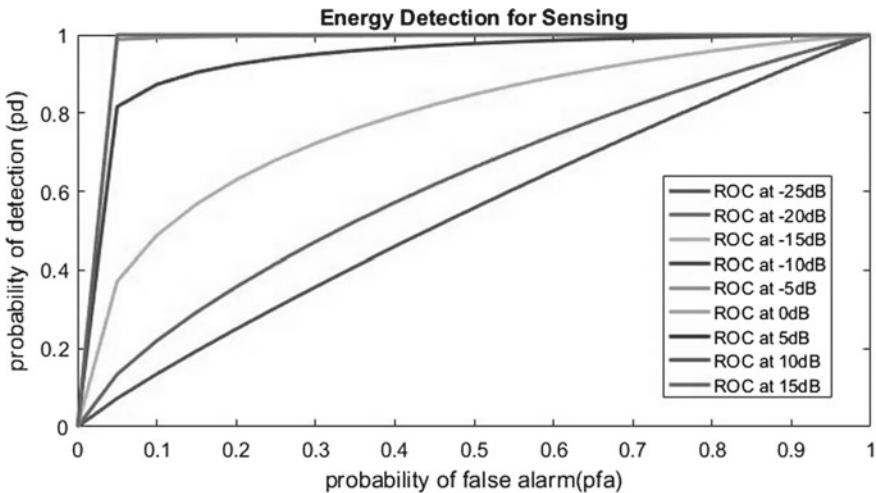


Fig. 4 ROC between probability of detection and probability of false alarm of Energy Detection sensing for SNR from -25 to 5 dB for $K = 10^5$ iterations

References

1. Ramani, V., Sharma S.K.: Cognitive radios: a survey on spectrum sensing, security and spectrum handoff. *China Commun.* **14**(11), 185–208 (2017)
2. Ghosh, G., Das, P., Chatterjee, S.: Simulation and analysis of cognitive radio system using matlab. *Int. J. Next-Gener. Netw. (IJNGN)* **6**(2), 31 (2014)
3. Bixio, L., Ottonello, M., Raffetto, M., Regazzoni, C.S.: (EURASIP Member) Comparison among cognitive radio architectures for spectrum sensing. *EURASIP J. Wirel. Commun. Netw.* **2011**(11), 749891 (2011)
4. Ali, A., Hamouda, W.: Advances on spectrum sensing for cognitive radio networks: theory and applications. *IEEE Commun. Surv. Tutor.* **19**(2), 1277–1304 (2017)
5. Jayakrishna, P.S., Sudha, T.: Energy efficient wireless sensor network assisted spectrum sensing for cognitive radio network. In: 2017 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing, pp. 1–7. IEEE (2017)
6. Cosovic, I., Jondral, F.K., Buddhikot, M.M., Kohno, R.: Comparison among cognitive radio architectures for spectrum sensing. *EURASIP J. Wirel. Commun. Netw.* **2008** (2008)
7. Zeng, Y., Liang, Y. C., Hoang, A. T., Zhang, R.: A review on spectrum sensing for cognitive radio: challenges and solutions. *EURASIP J. Adv. Signal Process.* **2010**(1), 381465 (2010)
8. Zhang, D., Chen, Z., Ren, J., Zhang, N., Awad, M.K., Zhou, H., Shen, X.S.: Energy-harvesting-aided spectrum sensing and data transmission in heterogeneous cognitive radio sensor network. *IEEE Trans. Veh. Technol.* **66**(1), 831–843 (2017)
9. Cao, X., Zhou, X., Liu, L., Cheng, Y.: Energy-efficient spectrum sensing for cognitive radio enabled remote state estimation over wireless channels. *IEEE Trans. Wirel. Commun.* **14**(4), 2058–2071 (2015)
10. Ren, J., Zhang, Y., Ye, Q., Yang, K., Zhang, K., Shen, X.S.: Exploiting secure and energy-efficient collaborative spectrum sensing for cognitive radio sensor networks. *IEEE Trans. Wirel. Commun.* **15**(10), 6813–6827 (2016)
11. Fahimi, M., Ghasemi, A.: A distributed learning automata scheme for spectrum management in self-organized cognitive radio network. *IEEE Trans. Mob. Comput.* **16**(6), 1490–1501 (2017)

12. Nguyen, V.-D., Shin, O.-S.: Cooperative prediction-and-sensing-based spectrum sharing in cognitive radio networks. *IEEE Trans. Cogn. Commun. Netw.* **4**(1), 108–120 (2018)
13. Qiu, L., Liu, S., Zhang, Y., Zhu, Y., Tang, K., Zheng, Y.: A 0.9–2.6 GHz cognitive radio receiver with spread spectrum frequency synthesizer for spectrum sensing. *IEEE Sens. J.* **17**(22), 7569–7577 (2017)
14. Zhao, G., Ma, J., Li, G.Y., Wu, T., Kwon, Y., Soong, A., Yang, C.: Spatial spectrum holes for cognitive radio with relay-assisted directional transmission. *IEEE Trans. Wirel. Commun.* **8**(10), 5270–5279 (2009)

Design and Development of IoT-Based Framework for Indian Aquaculture



Zeenat Shareef and S. R. N. Reddy

Abstract Aquaculture, in addition to agriculture is one of the most sought after occupations for the people of coastal regions of India. These farmers depend on the yield generated from aquaculture farming for their livelihood. However, due to various environmental as well as social factors, the water quality in these farms deteriorates and leads to untimely death of fishes. Such a scenario leads to heavy losses to the farmers. These farmers follow traditional approaches in measuring water quality which is time-consuming and inefficient. They are unable to get any warnings to take measures to save their fishes from getting deteriorated. This paper presents design and development of an Internet of Things(IoT)-based framework which measures water quality in aquaculture farms and provides alerts to the farmers so that they can take necessary precautions and save themselves from hefty losses.

Keywords Aquaculture · Internet of things · Mobile application · Water quality · Wireless sensor network

1 Introduction

Aquaculture is cultivation of fishes, shrimps and other marine creatures under controlled environmental conditions. Aquaculture is one of the most sought after occupation for the people of coastal regions in India. Apart from meeting the domestic needs of the people, more than 14.5 million farmers depend on the fishery activities for their livelihood and foreign exchange earnings [6]. A lot of research is being

Z. Shareef (✉)

Department of Electronics and Communication Engineering, Indira Gandhi Delhi Technical University for Women, Delhi, India
e-mail: zeenat.taneez@gmail.com

S. R. N. Reddy

Department of Computer Science and Engineering, Indira Gandhi Delhi Technical University for Women, Delhi, India
e-mail: rammallik@yahoo.com

conducted on sustainability of Aquaculture in India by Indian as well as foreign universities [1–5]. It has been found that India contributes 2.5% of the world fish trade, which is only next to rice (10.4%), tea (16.4%), animal diet (4.3%). Thus, aquaculture is not only an important sector for the farmers of coastal regions but also for the economy of the country.

Sometimes, due to excessive pollution, human intervention, agricultural wastes, and other environmental reasons, the farm water gets polluted. There is a lot of stress on water and it is affected by a variety of contaminants and pathogens [7]. This changes the physiochemical parameters of water. As a result, the fishes and shrimps get affected and eventually die. Therefore, there is a need to measure the quality of water at regular intervals. Traditional method of measuring water quality by Indian farmers includes collecting water samples from farm and getting it tested in laboratory [8]. However, such a method is very lengthy, cumbersome and inefficient. The water parameters get changed over time during transport from farm to laboratory and the farmers are unable to get any warning to take precautions to save their fishes.

The penetration of information and communication technologies into various domains has solved various problems encountered in urban and rural areas [9, 10]. This paper presents an IoT-based framework which remotely monitors the water quality in aquaculture farms and provides alerts to the farmers. Section 2 describes related work in this domain. Section 3 presents the design and development of IoT-based Framework for Indian Aquaculture. Section 4 explicates results and future work.

2 Related Work

Incorporation of Internet of Things in various domains is resulting in creation of smart devices and hence smart spaces. In IoT, several resource constraint devices may interact with each other without human intervention which has led real-time monitoring of remote areas possible [11]. Incorporation of IoT into aquaculture domain can lead to real-time remote monitoring of water parameters. This section summarizes the work carried out by various researchers in the field of aquaculture.

Vijayakumar et al. [12], designed a water quality monitoring system using Raspberry Pi platform which is based on linux kernel. The authors measured water temperature, turbidity, pH, dissolved oxygen and conductivity. These water parameters were transmitted to the cloud platform called IoT Bridge server through the Wi-Fi module. Yu Zhai et al. [13], designed dissolved oxygen sensor based on C8051F040 computing platform. The data was transmitted to the PC via UART communication using SRWF1028 module. The measured readings were compared with the iodometry tests and the results were found to be satisfactory.

Masakazu Arima et al. [14], developed an autonomous surface station for under-water passive acoustic observation of marine mammals. It was a solar-powered vehicle for environment monitoring and underwater acoustic monitoring of marine animals. It had an embedded device and 3 axis accelerometer for navigation. Lambrou

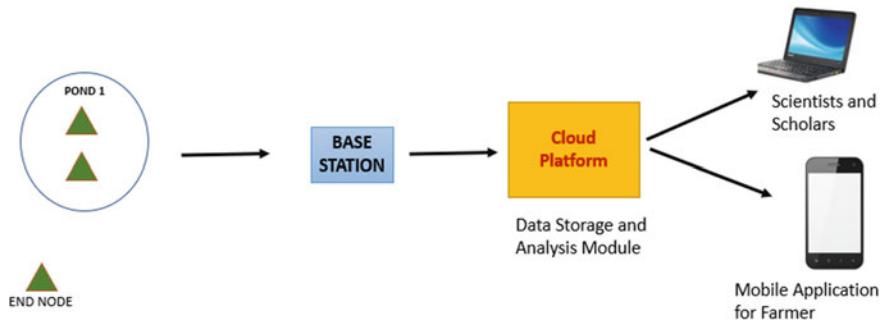


Fig. 1 Proposed IoT based framework for Indian aquaculture

et al. [15], designed a low-cost system for real-time monitoring and assessment of potable water quality at consumer sites. The water parameters that were measured include temperature, conductivity, pH, ORP, and turbidity. The sensor data was collected and sent to the remote station via Zigbee communication. The readings were further sent to the server through the ethernet.

Each author designed the system depending on the feasibility and requirements pertaining to a particular aquaculture field. The IoT framework designed for one field might not be suitable for another area. Hence, in this paper, an IoT-based Framework for Indian Aquaculture farmers of Western Godavari Region has been designed after extensive review, surveys and feedbacks and personal interviews.

3 System Design and Development

An extensive survey was conducted with fifty aquaculture farmers of western Godavari region to get an insight into their current farming habits, their problems, and their requirements. Such a survey is important for feasibility and acceptance of the system design by the end users (aquaculture farmers). The inference gathered from such a survey is: Requirement of a floating node to measure water parameters (water pH, water temperature, and dissolved oxygen) continuously and provide alert to end users through an Android Mobile Application.

In order to fulfill the requirements of the aquaculture farmers to measure the water quality, a framework was designed consisting of self-autonomous mobile sensor node to measure water parameters, a base station to receive measured water parameters from sensor node and transmit them to data storage and analysis module, a data storage, and analysis module for processing the data, and an Android Mobile Application for providing data and alerts to the end users. Figure 1 depicts the block diagram of the IoT-based Framework for Indian Aquaculture Farmers of Western Godavari Region.

Table 1 Sensors description

Sensor	Importance	Analog/Digital	Range
PT100 temperature sensor	Controls solubility of gases and other chemical reactions in water	Analog	0*-100* C
Galvanic dissolved oxygen	Important for support of life and waste decomposition	Analog	0–20 mg/L
Glass electrode pH sensor	Imbalance may make shrimp shell soft and endanger survival	Analog	0–14

3.1 Design of Self-autonomous Sensor Node

Considering the environmental conditions prevailing in the aquaculture farms and the requirement of the end users, a solar-powered sensor node was designed. The node is called self-autonomous because there is no requirement of human intervention. The node was powered by a 10 W/12 V solar panel and a 12 V/7.2 Ah rechargeable battery. Once the node is turned on, it automatically measures water parameters and transmits the data to the base station.

The self-autonomous sensor node is based on Intel Edison platform. Intel Edison is a computer on chip designed by Intel as a development system for Internet of Things devices. The Intel Edison SoC is a 22 nm Intel Atom “Tangier” (Z34XX) that includes two Atom Silvermont cores running at 500 MHz (running Yocto Linux) and one Intel Quark core at 100 MHz(running ViperOS). The self-autonomous sensor node measure water parameters namely, dissolved oxygen, water temperature, and water pH. Water quality measuring sensors along with their signal conditioning circuit were interfaced with the analog pins of Intel Edison. The accuracy of the sensors was improved through smoothening to remove overlays and anomalies. Table 1 presents details of the sensors. The measured data was transmitted from the sensor node to the base station through Zigbee communication module. Figure 2 depicts block diagram of self-autonomous sensor node.

3.2 Base Station

The base station collects the water parameter data from the self-autonomous sensor nodes and transmits them to the data storage and analysis module through Wi-Fi. The base station is also based on Intel Edison platform. Intel Edison has built-in Wi-Fi module which eliminates the need of interfacing any external hardware, thus reducing on cost and size of the base station. The integrated Wi-Fi module works on dual band (2.4 and 5 GHz) 802.11 a/b/g/n standard. AES is present in WLAN hardware for faster data encryption. The WLAN subsystem provides Wi-Fi protected

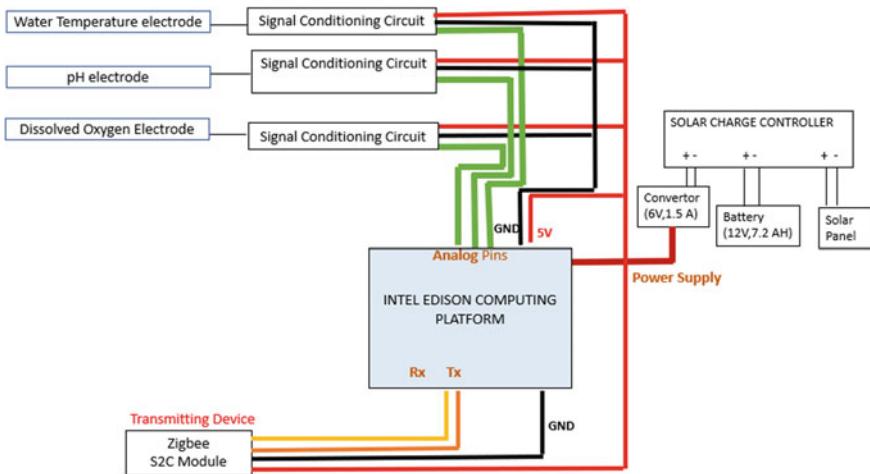


Fig. 2 Block diagram of self-autonomous sensor node

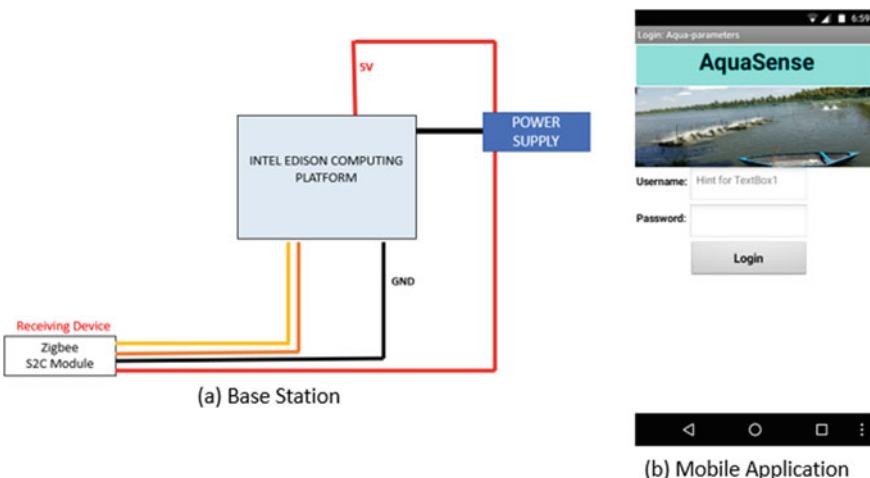


Fig. 3 a Base station design b Login page of customized android application

setup (WPS). Figure 3a presents the block diagram of Base Station, which may be powered as per the convenience of the farmer.

3.3 Data Storage and Analysis Module

This module is based on cloud platform. It can aggregate the data into the channel, visualize the data in graphical format and analyze the data if required. Each data is

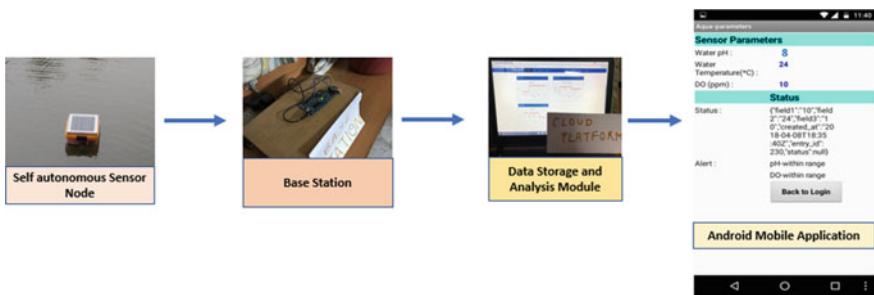


Fig. 4 Pilot deployment framework

stored with timestamp and date and is assigned a unique entry id. One can retrieve the data and perform analysis on the data if required. The data storage and analysis module may be placed at a fishery department of the area, where, data from different ponds could be stored for analysis by scientists or research scholars in the department.

3.4 An Android Mobile Application

A customized Android Mobile Application has been developed for the end users where they could visualize the data as well as get alerts whenever required. The mobile application has a login page and after successful login, it depicts the data and alert to the aquaculture farmers. Figure 3b depicts log in page of mobile application.

4 Result and Future Work

A pilot deployment of the system was carried out in Agra Canal, Okhla to test system parameters. It was observed that the system performed well. At present, the accuracy of the sensors is 87% for water temperature sensor, 89% for pH sensor, and 80% for dissolved oxygen sensor. There were issues with the placement of the sensor node in pond water, however, it was stabilized with the help of bricks and stones. Figure 4 depicts the pilot deployment of the system.

In future, we plan to expand the number of nodes to create a mesh network and deploy the system in Western Godavari region and collect feedback from the aquaculture farmers regarding the system.

Acknowledgements We would like to thank our collaborators at SRKR Engineering College, Bhimavaram for providing us with initial support for performing survey and case studies. We have all necessary permissions from participants for publishing this study. We would also like to extend our heartfelt thanks to Intel and DIC [MHRD] for providing prototyping platform to carry out our experiments.

References

1. Bagchi, A., Jha, P.: Fish and fisheries in indian heritage and development of pisciculture in India. *Rev. Fish. Sci.* **19**(2), 85–118 (2011)
2. Jana, B., Jana, S.: The potential and sustainability of aquaculture in India. *J. Appl. Aquac.* **13**(3–4), 283–316 (2003)
3. Valderrama, D., Iyemperumal, S., Krishnan, M.: Building consensus for sustainable development in aquaculture: a delphi study of better management practices for shrimp farming in India. *Aquac. Econ. Manag.* **18**(4), 369–394 (2014)
4. Gopakuma, K.: Indian aquaculture. *J. Appl. Aquac.* **13**(1–2), 1–10 (2003) (Tailor & Francis)
5. Gupta, N., Bower, S.D., Raghavan, R., Danylchuk, A.J., Cooke, S.J.: Status of recreational fisheries in India: development, issues, and opportunities. *Rev. Fish. Sci. Aquac.* **23**(3), 291–301 (2015)
6. Fisheries and aquaculture department. http://www.fao.org/fishery/countrysector/naso_india/en
7. Penmetsa, A.R.K.R., Muppudi, S., Popuri, R., Golla, S.B., Tenneti, R.: Aquaculture and its impact on ground water in East Godavari District Andhra Pradesh, India—a case study. *Int. Res. J. Environ. Sci.* **2**(10), 101–106 (2013). ISSN 2319–1414
8. Penmetsa, A., Muppudi, S.R., Popuri, R., Golla, S.B., Tenneti, R.: Impact of aquaculture on physico-chemical characteristics of water and soils in the coastal tracts of East and West Godavari Districts, Andhra Pradesh, India. *Int. J. Eng. Trends Technol. (IJETT)* **6**(6) (2013)
9. Churkina, G.: Modeling the carbon cycle of urban systems. *Ecol. Model.* **216**, 107–113 (2008)
10. Silvani, X., Morandini, F., Innocenti, E.: Evaluation of a wireless sensor network with low cost and low energy consumption for fire detection and monitoring. Springer Science + Business Media New York, Manufactured in the United States (2014). <https://doi.org/10.1007/s10694-014-0439-9>
11. Hussain, M.I.: CSIT **5**, 87 (2017). <https://doi.org/10.1007/s40012-016-0136-6>
12. Vijayakumar, N., Ramya, R.: The real time monitoring of water quality in IoT environment. In: IEEE Sponsored 2nd International Conference in Information, Embedded and Communication Systems (2015)
13. Zhao, Yu., Sun, L., Li, M.-F.: The research about detection of dissolved oxygen in water based on C8051F040. In: International Conference on Information Engineering and Computer Science (2009). <https://doi.org/10.1109/iciecs.2009.5363610>
14. Arima, M., Asahi, T.: Development of an autonomous surface station for underwater passive acoustic observation of marine mammals. In: OCEANS 2016–Shanghai
15. Theofanis, L., Panayiotou, G., Anastasiou, C.: A low-cost system for real time monitoring and assessment of potable water quality at consumer sites. In: 2012 IEEE Sensors, pp. 1–4 (2012)

Top- $k\%$ Concept Stratagem for Classifying Semantic Web Services



Aradhana Negi and Parminder Kaur

Abstract Top- k processing methodology is very popular among query processing in relational databases. The high influence of Top- k processing has been manifested in numerous application domains and database-related research areas. In this paper, the Top- k processing methodology has been adopted for the classification of Semantic Web Services (SWSs). It introduces the definition of the foundational unit of the Concept-sense Knowledge Base (CSKB) and Top- $k\%$ concept stratagem for classifying services to predefined categories in CSKB. The Top- $k\%$ concept scheme is implemented on OWLS-TC V₄ dataset. The outcomes of various performed experiments not only justify the implications of the introduced notion but also reveal the efficacy of classification time.

Keywords Concept-sense knowledge base (CSKB) · Service classification · Service discovery · Top- $k\%$ · Semantic web services (SWSs)

1 Introduction

Today's information search systems have a huge amount of data to search the potential answer for a query. Several types of information retrieval techniques [1] are being exploited to search the query results. Two most crucial aspects, which are encouraging either the exploitation of existing or development of new searching algorithms are: (1) rapid response time and (2) quality of results. Both of these aspects are the pillar of any successful information retrieval mechanism. Toward the effectiveness of these pillars, Top- k processing methodology [2] has achieved success in query optimization, indexing and ranking scheme, query processing, and document retrieval. The Top- k processing methodology aims to “*compute the results with highest scores first so that execution can terminate earlier after the top K results have been generated*” [3]. The score function always varies according to the applications practicality

A. Negi (✉) · P. Kaur

Department of Computer Science, Guru Nanak Dev University, Amritsar, India
e-mail: aradhana.csersh@gndu.ac.in

Table 1 Top- k processing methodology-related work

Sr. No	Approach	Year	Dataset
1	Top- k keyword search for supporting semantics in relational databases using BA (blocking algorithm) and EBA (early-stopping blocking algorithm) [9]	2008	$\times \times \times \times$
2	A survey of Top- k query processing techniques using query models, data access methods, implementation levels, data and query certainty, and supported scoring functions [2]	2008	Survey
3	Effective Top- k keyword search in relational databases considering query semantics [10]	2009	DBLP
4	Combination of Top- k processing and semantic pruning in join-based algorithms for keyword search efficiency in XML databases [3]	2010	DBLP and XMark
5	An effective Top- k keyword search algorithm based on classified Steiner tree [11]	2012	DBLP
6	Scalable continual Top- k keyword search in relational databases [12]	2013	DBLP
7	Top- k keyword search with recursive semantics in relational databases [13]	2017	$\times \times \times \times$

requirements. This paper presents the implications of Top- k processing in SWSs classification.

SWSs have been emerged with convergence of Service-Oriented Computing and Semantic Web [4]. The objective of SWSs is to encourage machine-readable annotations with service description, using some ontology [5], for realizing the full automation of web service life cycle. The competent service classification in the service repository assists the service provider to publish their service at one side and ease the service consumer to discover the required services at another side. In service classification process, the web services with similar functionality are classified into the same category [6]. A Classified SWSs Knowledge base may have many service categories, e.g., NAICS (North American Industry Classification System) [7] has many predefined categories for registering non-semantic, i.e., Web Service Description Language (WSDL) while the publically available OWLS-TC V4 dataset has nine categories for SWSs.

The motivational scenario presented next, drive us to use Top- k processing methodology in SWSs classification. The successful results gained from initial experiments done in [8] have encouraged us to use Top- $k\%$ concept stratagem in classification. We have also observed that top 35–50% concepts in each category can easily describe the domain specialization.

Example Scenario: *If a service repository has ten categories like financial, tourism, art, etc., and a service provider wants to publish a restaurant-related service in that repository, then as per existing approaches, the service will match with every*

Table 2 Web service classification-related work

Sr. No	Approach	Year	Dataset
1	Error correcting output codes with logistic model trees-based classifier [14]	2018	OWLS- TC ₄
2	Mathematical based model of naive Bayes classifier [15]	2016	OWLS- TC ₃
3	LDA-SVM active learning for scalable service classification [16]	2016	WS-DREAM
4	Omiotis measure of semantic relatedness with SVM and k-nearest neighbor classifier [17]	2016	OWLS- TC ₂
5	Construction of service interface graph for composable services [18]	2016	OWLS- TC ₄
6	Text mining approach for WSDL based service classification using 600 web services of 8 categories [19]	2015	Self-composed
7	SVM, naïve Bayes, decision tree, and BPNN for WSDL-based service classification [6]	2015	WSDL
8	Naïve Bayes, SVM, and REP tree to classify filtered 951 WSDL files (from Internet) [20]	2012	WSDL
9	Rank services with set quality attributes using naïve based, Markov blanket and tabu search [21]	2012	QWS dataset
10	Interface description classification using binarized and multiclass SVM, logistic regression, perceptron, and passive-aggressive [22]	2011	WSDL files from [23]
11	Ontology-based web service classification for OWL-S [24]	2011	× × × ×
12	Concept ranking with semantic measures [25]	2011	OWLS- TC ₃

service of predefined categories in the repository to match and store it in a category, which in our opinion is a time-consuming process. Now, let's say the repository can store services at their concept level and further those service concepts can be ranked using some ranking function and further for the service, to be published, if we can allow the match-making process to match initially with let's say top 40–50% concepts in each category then this approach can return matching results earlier than the results obtained after matching with all services of each category and can broadly classify the service to a category.

To attest the motivational scenario, this paper has emphasized on a sequence of experimentation. The remaining paper is structured into five sections. In Sect. 2, the research work related to Top- k processing methodology and web services classification is presented. In Sect. 3, the proposed Top- $k\%$ concept stratagem is described using definitions. Next, the proposed stratagem is evaluated and the results are presented in Sect. 4. Lastly, Sect. 5 gives the concluding remarks.

2 Related Work

In this section, an overview of the related work is presented in two parts: (1) Top- k processing methodology-related work and (2) Existing approaches for web service classification. Table 1 presents the Top- k processing methodology in relational databases varying from handling semantic in query to search XML documents and is confined to the work published from the year 2008 to 2017. Table 2 shows the web service classification approaches proposed in 2011–2018 in reverse chronological order. From Table 2, it is clear that 9 out of 12 research works signifies the use of machine learning techniques for service classification and also none of the research work has included more than 1 semantic formalism, i.e., WSMO, SAWSDL, etc., together with OWLS or WSDL. The cross sign ($\times \times \times \times$) in Tables 1 and 2 depicts the lack of dataset information.

3 Proposed Top- $k\%$ Concept

To fulfill the objectives of hybrid approach-based generic framework presented in [8], this paper proposes the classification strategy for SWSs. The Concept-sense Knowledge Base (CSKb) is the mainstay of the hybrid approach-based generic framework. SWSs (all extracted terms) are stored in the predefined categories of CSKb as an Octuple (O_p). Services store together irrespective of their language format restrictions. In this paper, only the definition part which is required for Top- $k\%$ concept stratagem, i.e., Octuple (O_p) of CSKb is given.

Definition 1 The foundational unit of CSKb is an Octuple (O_p):

$$O_p = \{ c, pOs, s, cat, sf, cf, wt, r^{nk} \}$$

where c is a concept of service description and $c \in cat$, cat refers to the category/domain of a service, pOs is the part of speech of c (noun/verb/adjective), s is the respective sense of c (captured from WordNet), sf is a pointer to an array that stores the list of services containing c , cf is the frequency of c in a cat , wt is the associated weight of c in a cat , r^{nk} indicates the rank of c in cat .

According to [25], the rank of a \mathbf{c} depends upon three points: (1) number of services using \mathbf{c} in a \mathbf{cat} , (2) the presence and significance of parent concept of \mathbf{c} in \mathbf{cat} , and (3) the presence and significance of child concepts of \mathbf{c} in \mathbf{cat} . We have used a variant of [25] and included the list of synonym and sense to determine the presence and significance of equivalent words and sense for final rank evaluation. The context analysis, i.e., related sense and part-of-speech tagging of service description are captured using natural language processing techniques as in [8].

Definition 2 For the given set of octuplets $\mathbf{O}_{P_1}, \mathbf{O}_{P_2}, \mathbf{O}_{P_3}, \dots, \mathbf{O}_{P_n} \in \mathbf{O}_P$, the $\mathbf{r}^{\text{nk}}(\mathbf{O}_{P_n})$ of the n th octuplet is given as

$$\sum_{n=1}^k \mathbf{cf}_n * \mathbf{wt}_c + \mathbf{Sup}_n * \mathbf{wt}_{\text{sup}} + \mathbf{Sub}_n * \mathbf{wt}_{\text{sub}} + \mathbf{Syn}_n * \mathbf{wt}_{\text{syn}} + \mathbf{Sen}_n * \mathbf{wt}_{\text{sen}}$$

Here, the superordinate (\mathbf{Super}_n), subordinate (\mathbf{Sub}_n), equivalent words (\mathbf{Syn}_n), and sense (\mathbf{Sense}_n) reveals the importance of hypernym, hyponym, synonym, and sense that belongs to a service and is also present in category at the same time. The list of hypernym, hyponym, synonym, and sense of concept are prepared from WordNet.¹ The \mathbf{wt} specifies the weights of the respective terms. To calculate associated \mathbf{wt} in its category, a count function is applied in each category. This count function captures the total number of concepts present in a category at the time of calculation and serves a denominator to weight measurement. Formally, \mathbf{wt} is the division of \mathbf{cf} to the total category concepts. The total number of concept in a category updates during the addition of every new concepts in that domain. The updating CSKb leads to update the dependent elements such that rank and frequency of the concept.

Definition 3 For every $\mathbf{c} \in \mathbf{cat}$, the Top- $k\%$ concepts are the Top- $k\%$ Octuple $\mathbf{O}_{P_1}, \mathbf{O}_{P_2}, \mathbf{O}_{P_3}, \dots, \mathbf{O}_{P_n} \in \mathbf{O}_P$, which is ordered by \mathbf{r}^{nk} and $\mathbf{r}^{\text{nk}}(\mathbf{O}_{P_i}) \geq \mathbf{r}^{\text{nk}}(\mathbf{O}_{P_{i+1}}) \geq \mathbf{r}^{\text{nk}}(\mathbf{O}_{P_{i+2}}) \geq \dots \geq \mathbf{r}^{\text{nk}}(\mathbf{O}_{P_k})$, here, i is the top concept of the category.

We have used the Multiclass Support Vector Machine (MSVM) [26] for implementing classification strategy. MSVM is a widely used supervised learning-based classifier, which builds a separate hyperplane with the maximal margin for available classes. According to the MSVM, if each service $\mathbf{n} = \{S_1, S_2, \dots, S_n\} \in \mathbf{S}$ belongs to a category \mathbf{cat}_j in $\mathbf{t} = \{\mathbf{cat}_1, \mathbf{cat}_2, \dots, \mathbf{cat}_t\} \in \mathbf{cat}$, then the feature space $\mathbf{fs} = \{\mathbf{fs}_1, \mathbf{fs}_2, \dots, \mathbf{fs}_n\}$ contains all the unique terms of \mathbf{n} services and each S_i is mapped to the feature vector $\mathbf{fv} = \{\mathbf{fv}_1, \mathbf{fv}_2, \dots, \mathbf{fv}_n\}$ such that \mathbf{fv} denotes the presence or absence (1 or 0) of feature vector in service S_i . All the mapped featured vectors of services are passed through the MSVM classifier to identify the corresponding category of the services.

¹<http://wordnetweb.princeton.edu/perl/webwn>.

Table 3 OWLS dataset services statistics

Category	Total services	20% services	40% services	60% services	80% services
Education	285	57	114	171	228
Communication	58	12	23	35	46
Medical	53	11	22	32	42
Food	34	7	14	20	27
Travel	165	33	66	99	132
Geography	60	12	24	36	48
Economy	359	72	144	215	287
Simulation	16	3	6	10	13
Weapon	40	8	16	24	32

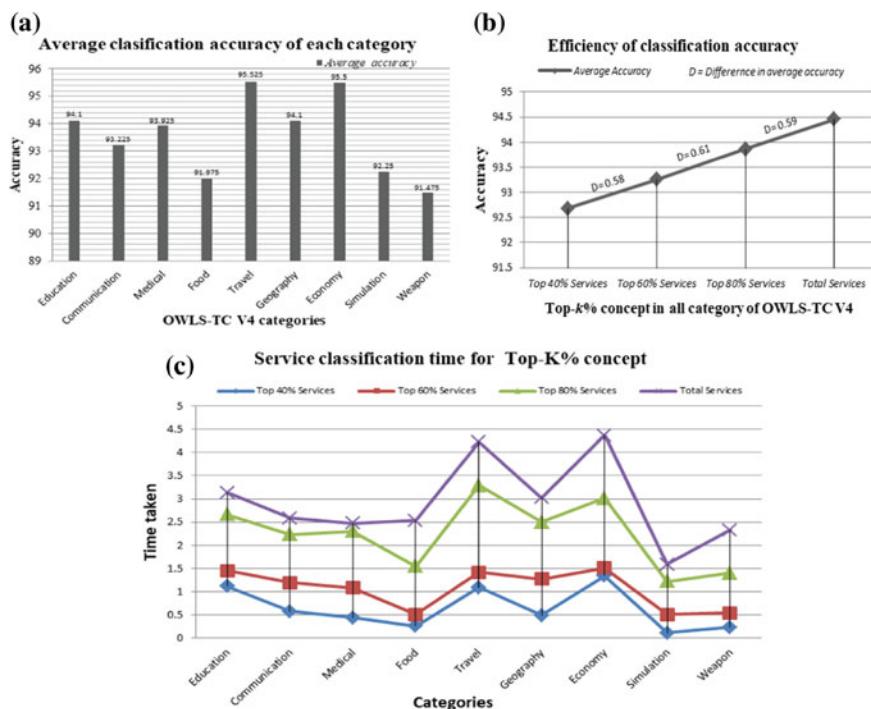
4 Implementation

To attest the proposed approach, the paper uses OWLS-TC V₄ dataset. We also planned to use extended private WSMO dataset [27], as it was used in [8], but unfortunately, this dataset is too small to classify using the MSVM. Hence, the experiments are curbed to OWLS-TC V₄ dataset that has 9 categories with 1083 services, 48 ontologies, and 43 queries. To train the classifier, the Spyder platform of Python on Windows10 with 16 GB RAM has been used. All the experimental observations are collected using tenfold cross validation. Table 3 describes the statistics of OWLS-TC V₄ dataset services. The paper has examined classification accuracy at different Top- $k\%$ concept levels such as Top-40%, Top-60%, Top-80%, 100%.

During the classification accuracy investigation, we have observed that accuracy varies from 91.3 to 97.4% (see Table 4) for all the considered four Top- $k\%$ concept levels. Figure 1a shows the achieved average classification accuracy of all Top- $k\%$ concept level together. From these obtained accuracies, we have observed (see Fig. 1b) that an average increment of 0.59% remains at every level from Top-40 to 100% concepts. This little variation of accuracies exists due to the coverage of concepts. It is also observed that the results for Top-80% are very close to the results for 100% in a very time efficient manner as given in Fig. 1c. This implies that instead of considering all services in each category for classification, we can store services as concepts and classify them with an average accuracy level of 93.86% at Top-80% concept level also. In the earlier approaches, the services, to be published, are classified by matching them with every service in each predefined category which took a high amount of time to present final results, whereas our proposed approach outperforms [17, 18] in terms of accuracy and time efficiency for service classification by allowing to match with Top- $k\%$ concepts from each category.

Table 4 Classification accuracy percentage at different Top- $k\%$ concept levels

Category	Top 40% services	Top 60% services	Top 80% services	Total services
Education	93.4	93	94.8	95.2
Communication	92.5	93.1	93.5	93.8
Medical	92.1	93.7	94.2	95.7
Food	91.7	91.7	92.1	92.4
Travel	94.3	95.4	95.8	96.6
Geography	93.5	94	94.2	94.7
Economy	93.9	94.6	96.1	97.4
Simulation	91.4	92.2	92.7	92.7
Weapon	91.3	91.6	91.4	91.6

**Fig. 1** Average accuracy and time efficiency of Top- $k\%$ concept stratagem

5 Conclusion

This paper combines the Top- $k\%$ concept stratagem, inspired from relational databases, in semantic web services classification. The paper introduces the definitions of Concept-sense Knowledge Base (CSKB) and Top- $k\%$ concept stratagem for service classification. In experimental results, it is observed that applying the Top- $k\%$ concept stratagem on categories that stores services in the form of concepts are time efficient, whereas the accuracy of classification varies from 91.3 to 97.4%. The multiclass support vector machine-based classifier results conclude that to get timely results with a very little variation of accuracy, one can classify services using Top-40%–Top-80% concepts.

As the experiments of this paper are limited to OWLS-TC V4 but in future, we are trying to expand available WSMO services for classification along with other semantic formalisms. Also in the near future, we will exploit the other version of multiclass support vector machine and some other machine learning classifiers.

References

1. Büttcher, S., Clarke, C., Cormack, G.V.: *Information Retrieval: Implementing and Evaluating Search Engines*. MIT Press, Cambridge (2010)
2. Ilyas, F., Beskales, G., Soliman, M.A.: A survey of top-k query processing techniques in relational database systems. *ACM Comput. Surv.* **40**(4), 11:1–11:58 (2008)
3. Chen, L.J., Papakonstantinou, Y.: Supporting top-K keyword search in XML databases. In: *Proceedings of International Conference on Data Engineering*, pp. 689–700 (2010)
4. Jemima, D.D., Karpagam, G.R.: Conceptual framework for semantic web service composition. In: *2016 International Conference on Recent Trends in Information Technology (ICRTIT)*, pp. 1–6 (2016)
5. Otero-Cerdeira, L., Rodríguez-Martínez, F.J., Gómez-Rodríguez, A.: Ontology matching: a literature review. *Expert Syst. Appl.* **42**(2), 949–971 (2015)
6. Yang, J., Zhou, X.: Semi-automatic web service classification using machine learning. *Int. J. Sci. Technol.* **8**(4), 339–348 (2015)
7. Corella, M., Castells, P.: A heuristic approach to semantic web services classification. *Knowl.-Based Intell. Inf. Eng. Syst.* **4253**, 598–605 (2006)
8. Negi, A., Kaur, P.: Examination of sense significance in semantic web services discovery **771** (2019)
9. Wang, B., Yang, X.-C., Wang, G.-R.: Top-K keyword search for supporting semantics in relational databases. *Ruan Jian Xue Bao/J. Softw.* **19**(9), 2362–2375 (2008)
10. Xu, Y., Ishikawa, Y., Guan, J.: Effective Top-k keyword search in relational databases considering query semantics. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* **5731**, 172–184 (2009)
11. Yang, Y., Tang, M., Zhong, Y., Zhang, Z., Guo, L.: An Effective top-k keyword search algorithm based on classified steiner tree. *Web-Age Inf. Manag.*, 276–288 (2012)
12. Xu, Y., Guan, J., Li, F., Zhou, S.: Scalable continual top-k keyword search in relational databases. *Data Knowl. Eng.* **86**, 206–223 (2013)
13. Liu, D., Liu, G., Zhao, W., Hou, Y.: Top-k keyword search with recursive semantics in relational databases. *Int. J. Comput. Sci. Eng.* **14**(4), 359 (2017)
14. Mavridou, E., Giannoutakis, K.M., Kehagias, D., Tzovaras, D., Hassapis, G.: Automatic categorization of web service elements. *Int. J. Web Inf. Syst.* **14**(2), 233–258 (2018)

15. Liu, J., Tian, Z., Liu, P., Jiang, J., Li, Z.: An approach of semantic web service classification based on naive bayes. In: 2016 IEEE International Conference on Services Computing (SCC), pp. 356–362 (2016)
16. Liu, X., Agarwal, S., Ding, C., Yu, Q.: An LDA-SVM active learning framework for web service classification. In: 2016 IEEE International Conference on Web Services (ICWS), pp. 49–56 (2016)
17. Sharma, S., Lather, J.S., Dave, M., Sharma, B.S.: Semantic approach for web service classification using machine learning and measures of semantic relatedness. *Serv. Oriented Comput. Appl.* **10**(3), 221–231 (2016)
18. Kamath, S.S., Ananthanarayana, V.S.: Semantics-based web service classification using morphological analysis and ensemble learning techniques. *Int. J. Data Sci. Anal.* **2**(1–2), 61–74 (2016)
19. Nisa, R., Qamar, U.: A text mining based approach for web service classification. *Inf. Syst. E-bus. Manag.* **13**(4), 751–768 (2015)
20. Yuan-jie, L., Jian, C.: Web service classification based on automatic semantic annotation and ensemble learning. In: 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops and PhD Forum, pp. 2274–2279 (2012)
21. Mohanty, R.: Classification of web services using bayesian network. *J. Softw. Eng. Appl.* **05**(04), 291–296 (2012)
22. Bennaceur, A., Johansson, R., Moschitti, A., Sykes, D., Spalazzese, R.: Machine learning for automatic classification of web service interface descriptions. *Leveraging Appl. Form. Methods, Verif. Valid.*, 220–231 (2012)
23. Heß, A., Kushmerick, N.: Learning to Attach Semantic Metadata to Web Services, pp. 258–273. Springer, Berlin (2003)
24. Wu, H., Guo, C.: The research and implementation of web service classification and discovery based on semantic. *Int. Conf. Comput. Support. Coop. Work Des.*, 381–385 (2011)
25. Farrag, T.A., Saleh, A.I., Ali, H.A.: ASWSC: automatic semantic web services classifier based on semantic relations. *Int. Conf. Comput. Eng. Syst.*, 283–288 (2011)
26. Wang, Z., Xue, X.: Multi-class support vector machine. *Support Vector Machines Applications*, pp. 23–48. Springer International Publishing, Cham (2014)
27. Sangers, J., Frasincar, F., Hogenboom, F., Chepegin, V.: Semantic Web service discovery using natural language processing techniques. *Expert Syst. Appl.* **40**(11), 4660–4671 (2013)

Performance Investigation of Fractional PI Controller in Shunt Active Filter for a Three-Phase Three-Wire System



Ritu Sharma, Naveen Goel, Saji Chacko and Ravi Prakash Mahobia

Abstract Due to advancement in the properties of semiconductor devices, they are widely used in domestic as well as industrial application. These devices are showing nonlinear property and cause different power quality issues. Harmonic is one of the major power quality issues. To eliminate these harmonics, various technologies have been developed including passive filters, zigzag transformers, active filters, etc. Active filters can be used in different configurations such as series active filters, shunt active filters, etc. In this work, shunt active power filter has been used to eliminate the harmonics from the system and is based on the instantaneous active and reactive power theory. Pulse-width modulation current control technique has been used to generate pulses for voltage-source converter. PI controller is used to regulate the DC voltage. To obtain the accurate output result, tuning of PI controller parameters (proportionality constant K_p , integral constant K_i) is required. Fractional PI controller has been studied in this paper and its performance has been compared with the performance of conventional PI controller. The Use of fractional-order PI controller increases with the passage of time because its performance is better than the conventional PI controllers. It contains one more constant term, called as fractional constant with proportionality constant and integral constant for the tuning which provides accurate result.

Keywords Harmonic · Semiconductor devices · Passive filters · Active filters · Fractional PI controller · Conventional PI controller

R. Sharma (✉) · N. Goel · R. P. Mahobia

Department of Electrical and Electronics Engineering, SSTC, Junwani, Bhilai, Chhattisgarh, India
e-mail: ritusharma10893@gmail.com

S. Chacko

Electrical Engineering, Government Polytechnic, Durg, Chhattisgarh, India

1 Introduction

Three-phase distribution system is widely used to supply power to single-phase and three-phase loads. Since a few decades development and enhancement in properties of semiconductor devices, these devices fascinate the customers. These devices such as inverter-based home appliances, arc furnaces, adjustable-speed motor drives, etc., are used widely. These semiconductors show the nonlinear properties and cause various power quality issues, it may be unbalanced reactive power, voltage, current waveform distortion and harmonics, etc. Harmonics is the integer multiple of the fundamental frequency, and these components combine with the system quantity and distort the system current and voltage waveforms. Therefore, it is necessary to reduce these harmonics from the system.

Various methods are employed to reduce the harmonic current from the system such as using zigzag transformers, passive filters, active power filters (APF) [1, 2], hybrid filters. Hybrid filters can be formed by a combination of shunt passive filters with series active filters [3, 4] and also by combining series passive filters with series active filters [5]. Passive filters have different drawbacks such as occurrence of resonance, larger in size, etc. APFs has replaced the passive filters since few years. In this work, shunt active power filter (SAPF) [6] has been used to mitigate the harmonics. SAPFs injects or absorbs the compensating currents, equal to the harmonic current. And this compensating current added or subtracted with system current and cancels out the harmonic current from the system. SAPF can be divided into two topologies. First, includes voltage-source converters (VSC) [7, 8] and second includes current-source converters (CSC).

Basically, two types of converters is used for implementing the SAF, VSC, and CSC. The main function of the pulse-width modulation (PWM) controller is to enforce the converter to work as a controlled current source. The average energy exchange between SAF and the power system must be equal to zero. In VSCs, capacitor is used as the energy storage element while in CSC, inductors are used to store the energy. In this work, SAPF with VSC has been used.

PWM control technology has been used to control the switching pulses of the converter [9, 10]. PI controllers are used in the SAFs. It generates error signals and that error signals are fed back to the system. With the conventional proportional-integral (CPI) controller, the output result can be improved by varying the value of proportional constant K_p and integral constant K_i . While the fractional proportional integral (FPI) controller provides a better result as compared to CPI controller. FPI controller includes one more constant called as fractional constant λ with K_p and K_i for the tuning. In this work, conventional tuning has been done for CPI controller and FPI controller.

2 Instantaneous Reactive Power Theory

Transformation of quantities can be done in either frequency domain or in time domain. Various transformation theories have been developed such as instantaneous reactive power theory (p-q transformation theory) [11, 12], d-q transformation theory [13], etc. In this work, p-q transformation theory has been used, which defines a set of instantaneous power in the time domain. It uses Clarke's transformation approach.

$$\begin{bmatrix} v_\alpha \\ v_\beta \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} 1 & \frac{-1}{2} & \frac{-1}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{-\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} v_a \\ v_b \\ v_c \end{bmatrix} \quad (1)$$

Three-phase quantities are converted from a–b–c coordinates to α – β coordinates by using Eqs. (1) and (2). Equation (1) is used for source voltage transformation while Eq. (2) shows the load–current transformation into α – β coordinate.

$$\begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} 1 & \frac{-1}{2} & \frac{-1}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{-\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} i_{La} \\ i_{Lb} \\ i_{Lc} \end{bmatrix} \quad (2)$$

The advantages of p-q-0 transformation are that in α – β coordinate, the user has to deal with the magnitude of the quantity only and there is no need of phase consideration, therefore, complex calculations reduces to easy calculation.

The transformed voltage and current in α – β coordinate are used to calculate the active and reactive power in α – β coordinate by using Eq. (3).

$$\begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} v_\alpha & v_\beta \\ v_\beta & -v_\alpha \end{bmatrix} \begin{bmatrix} i_\alpha \\ i_\beta \end{bmatrix} \quad (3)$$

Real power and reactive power with voltage in α – β coordinate are used to calculate the compensating current by using Eq. (4). The compensating current obtained in α – β coordinate is again converted to a–b–c coordinate by using Eq. (5).

$$\begin{bmatrix} i_{C\alpha} \\ i_{C\beta} \end{bmatrix} = \frac{1}{v_\alpha^2 + v_\beta^2} \begin{bmatrix} v_\alpha & v_\beta \\ v_\beta & -v_\alpha \end{bmatrix} \begin{bmatrix} -\tilde{p} + \bar{p}_{loss} \\ -q \end{bmatrix} \quad (4)$$

$$\begin{bmatrix} i_{Ca} \\ i_{Cb} \\ i_{Cc} \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} 1 & 0 \\ \frac{-1}{2} & \frac{\sqrt{3}}{2} \\ \frac{-1}{2} & \frac{-\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} i_{C\alpha} \\ i_{C\beta} \end{bmatrix} \quad (5)$$

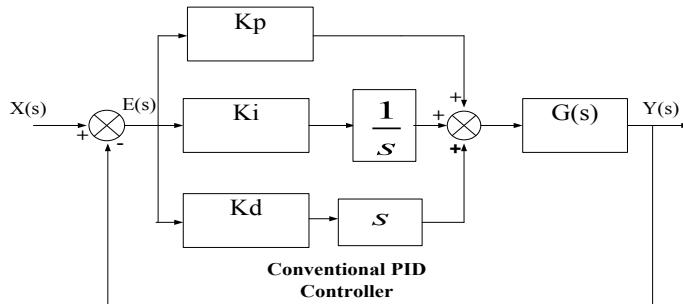


Fig. 1 Structure of conventional PID controller

3 Proposed Control Strategy

The controller designed for the three-phase three-wire system deals with the real power and reactive power of the system [14, 15]. In this work, a constant instantaneous power control strategy has been used. The SAF controller works in a closed-loop manner. It continuously senses the load current accordingly generates compensating current. This compensating current is used to generate pulses for the converter switches and cancel out the harmonic current from the system.

3.1 Conventional PI Controller

Controller parameters are adjusted by tuning the parameters of PID controller. Conventional PID controller has three parameters, i.e., proportionality constant K_p , integral constant K_i , and differential constant K_d as shown in Fig. 1. By varying the value of these constants, the required output can be obtained. PI controller is formed by equating differential constant to zero. For any system, the gain $G(s)$ for PI controller is given by Eq. (6).

$$G(s) = K_p + \frac{K_i}{s} \quad (6)$$

3.2 Fractional PI Controller

Presently, fractional-order PID controllers [16] are widely used in the science technology as a control of inverter, rectifier, boiler system, motors [17], etc. Fractional calculus mentioned in different books [18–21] is applied in FPI controllers. The structure of fractional PID controller is shown in Fig. 2. This fractional-order PID controller have five parameters K_p , K_i , K_d , and two fractional constants λ and μ .

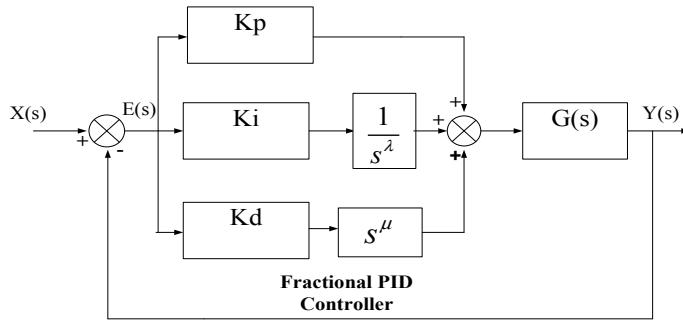


Fig. 2 Structure of fractional order PID controller

This controller is used as FPI controller by taking the value of K_d and μ equal to zero as shown in Eq. (7). FPI controllers give better result as compared to CPI controllers [16, 22].

$$G(s) = K_p + \frac{K_i}{s^\lambda} \quad (7)$$

where $G(s)$ = controlled plant, $X(s)$ = input to system, $Y(s)$ = output of the system, K_p and K_i are proportionality and integral constant, λ = fractional constant.

4 Methodology

Design of shunt active power filter includes mainly four parts:

1. Calculation of instantaneous active and reactive power,
2. Selection of power to be compensating,
3. Regulation of DC voltage,
4. Reference current calculation.

SAPF is formed by connecting all the blocks mentioned above [23]. Figure 3 shows the complete block diagram of all the components of SAPF, which is used to eliminate harmonics from three-phase three-wire systems. The system consists of nonlinear load, this nonlinear load contains the universal bridge of diode elements. SAPF has been connected at the point of common coupling between source and load. Coupling inductor has been used to reduce high di/dt .

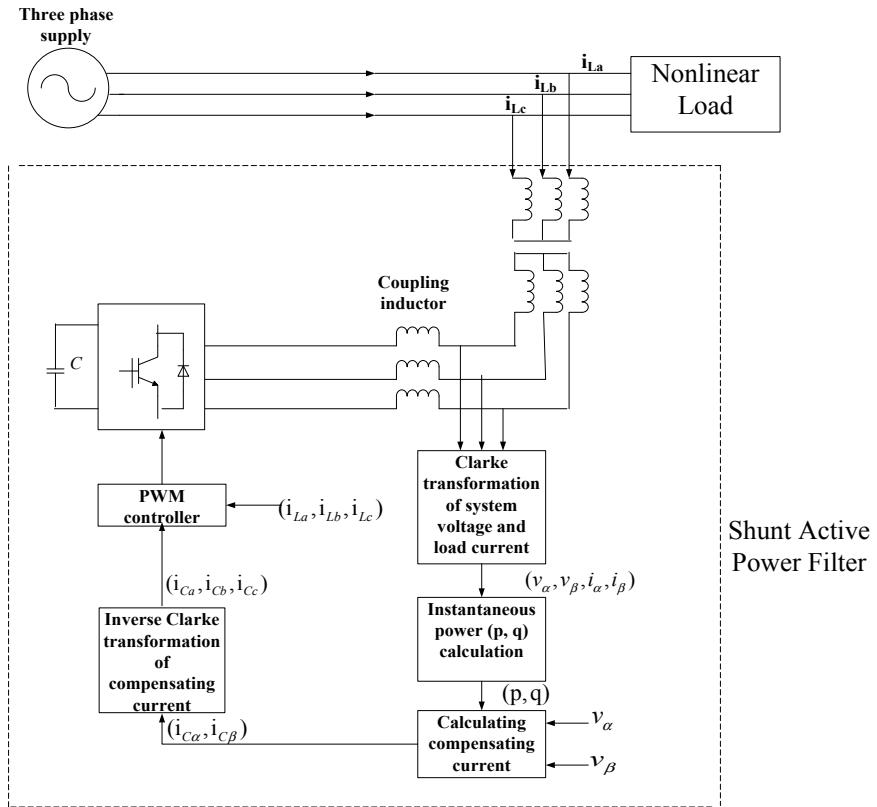


Fig. 3 Block diagram of shunt active power filter

Table 1 Simulation parameters

Supply voltage	440 V
System frequency	50 Hz
Coupling inductor	2 mH
DC bus voltage	850 V
PWM switching frequency	10 kHz
DC link capacitor	40 μ F

5 Results and Discussion

Performance of shunt active power filter has been done by using MATLAB Simulink. Table 1 shows the magnitude of all the parameters, which is used to design the Simulink model of shunt active power filter.

The simulation of the Simulink model has been done for 1 s. After simulation of the Simulink model of power system consisting of shunt active power filter results

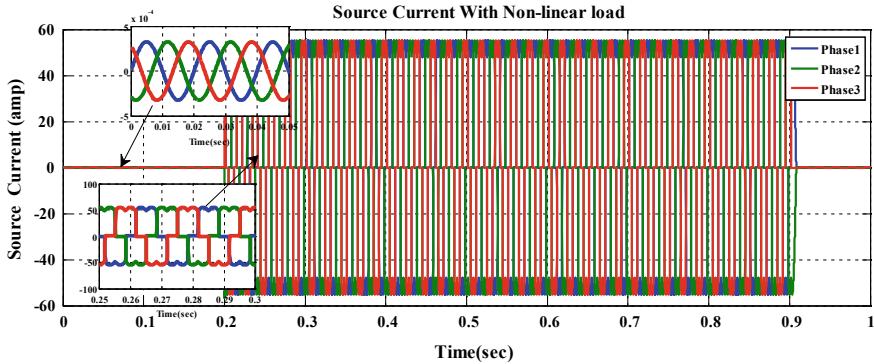


Fig. 4 System current waveform

has been obtained in the form of waveforms. Simulation process has been done for two conditions, these conditions are:

1. Output waveforms obtained for system with nonlinear load,
2. Output waveforms obtained for the system with nonlinear load and SAPF. It is also divided into two subparts:
 - 2.1. Waveforms for Shunt Active Power Filter with conventional PI controller,
 - 2.2. Waveforms for Shunt Active Power Filter with fractional PI controller.

5.1 System with Nonlinear Load

Linear loads in the system cause sinusoidal waveform of source voltage and source current. Nonlinear loads are used in domestic applications (microwave ovens, TVs, computers, etc.), and in industrial applications (arc furnaces, variable frequency drives, electronic ballasts, etc.) causes distortion in current waveform because of existence of harmonic component in the system. System-source voltage waveform does not contain any harmonic component, therefore, the voltage waveform is pure sinusoidal during complete simulation of 1 s. Each phase is 120° apart from each other.

The simulation results carried out for 1 s. From 0 to 0.2 s when there is only linear load is connected in system, the source current waveform is sinusoidal. At 0.2 s, a nonlinear load of universal bridge consist of diode element has been connected in the system and current waveform at supply side becomes non-sinusoidal resulting in generation of harmonic current components along with the fundamental component as shown in Fig. 4. This nonlinear load has been connected till 0.9 s.

FFT analysis of source current as shown in Fig. 5 shows that the system with nonlinear load injects harmonics to the system. Total harmonic distortion of any

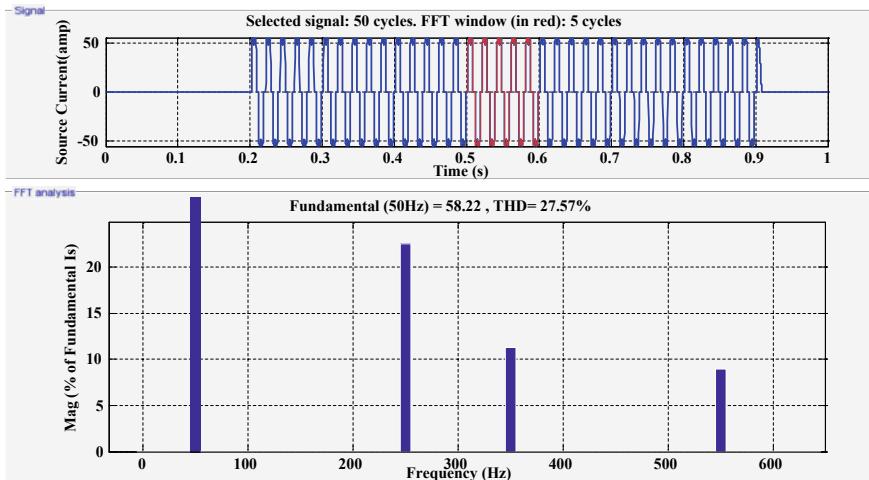


Fig. 5 FFT plot of source current

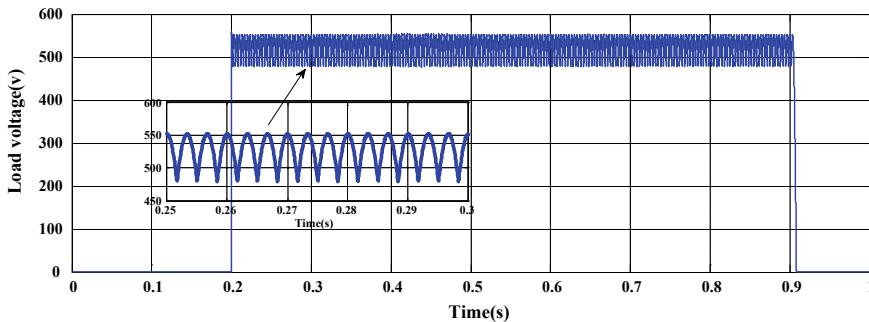


Fig. 6 DC voltage at load terminal

waveform shows the percentage of harmonic component present in the given signal. Here, THD of source current is 27.57%.

Nonlinear load connected across the system is three-phase-uncontrolled rectifiers of diode element, thus, voltage at the load terminal is pulsating DC voltage as shown in Fig. 6.

5.2 System with Nonlinear Load and Shunt Active Power Filter

At 0.3 s ,shunt active power filter has been connected to the system. The SAPF generates compensating current, which is equal to the magnitude of the harmonic

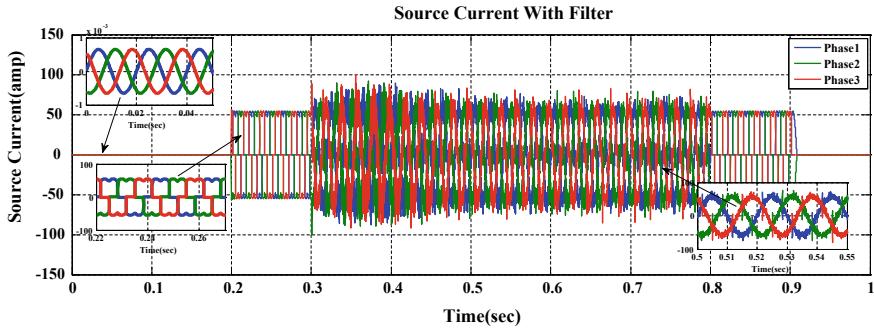


Fig. 7 Source current waveform after connecting SAPF

current. This compensating current cancel out the harmonics from the source current and distorted source current waveform again becomes sinusoidal as shown in Fig. 7.

SAPF uses CPI controller for the regulation of DC voltage by tuning the PI parameters. The capacitor DC voltage is compared with the fixed DC voltage known as reference DC voltage. The result obtained from PI controller is used as power loss in the converter circuit. In this work, FPI controller has been studied and its performance has been compared with the CPI controller. Thus, case II is categorized into two parts, as mentioned below.

5.2.1 SAPF with Conventional PI Controller

With conventional PI controller, shunt active power filter gives the following results:

The FFT analysis of source current has been done for five cycles as shown in Fig. 8, shows that the harmonics present in source current has been reduced from THD of 27.57% to 6.65% by using conventional PI controller.

5.2.2 SAPF with Fractional PI Controller

With fractional PI controller, shunt active power filter gives the following results:

The FFT analysis of source current as shown in Fig. 9 shows that the harmonics present in source current has been reduced from THD of 28.46% to 6.73% by using fractional PI controller. Five cycles of waveform is considered for the FFT analysis. 50 Hz has been taken as fundamental frequency.

The DC capacitor voltage for the system obtained for the conventional PI controller and fractional PI controller has been shown in Fig. 10.

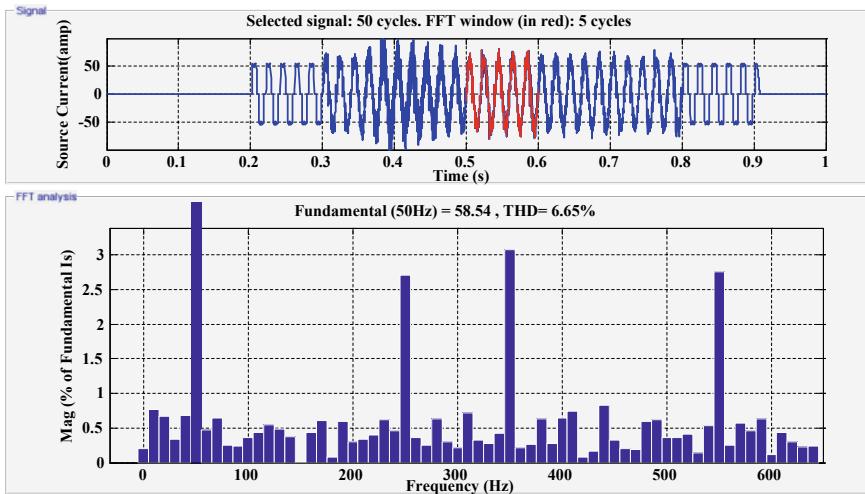


Fig. 8 FFT plot of source current after connecting SAPF with CPI controller

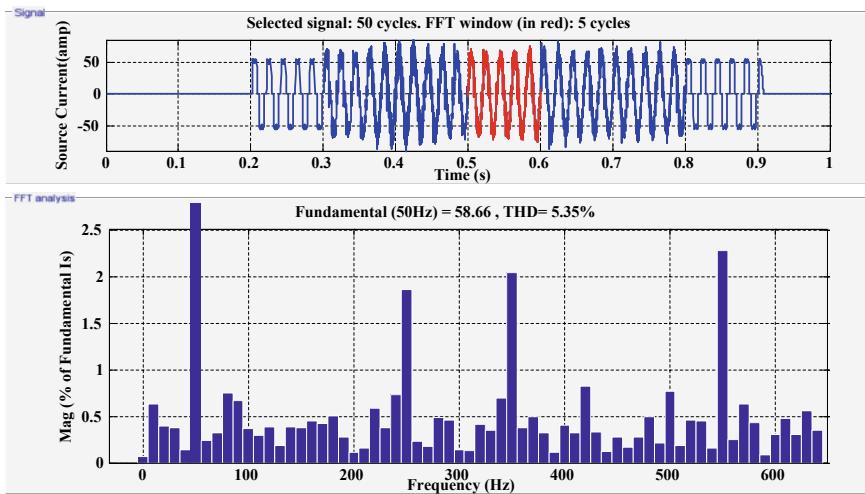


Fig. 9 FFT plot of source current after connecting SAPF with FPI controller

Comparison between PI and fractional PI controller regarding the DC voltage has been shown in Table 2. The peak overshoot in fractional PI controller is lower than the PI controller.

Table 3 shows that the use of CPI controller reduces harmonics from current THD of 27.57% to 6.65% while the FPI controller reduces current THD up to 5.35%. It shows that the performance of FPI controller is better than CPI controller.

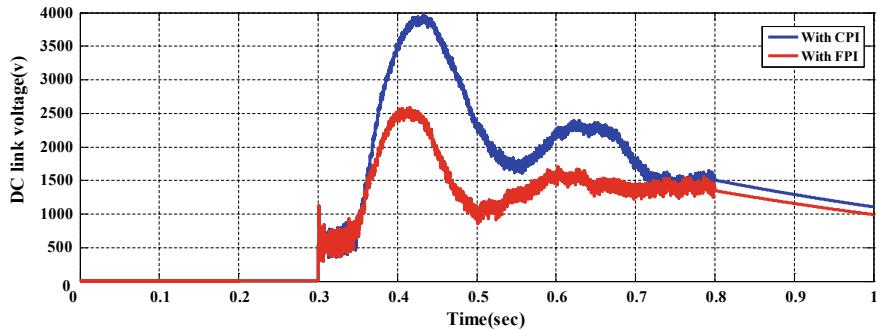


Fig. 10 DC capacitor voltage for CPI and FPI controller

Table 2 Comparison of CPI controller with FPI controller

Quantity	CPI controller	FPI controller
Rise time	8.96	10.22
Settling time	1.9e04	1.95e04
Overshoot	318.99	161.1
Peak	3.67e03	2.59e03
Peak time	8258	8310

Table 3 Comparison of current THD for CPI controller and FPI controller

Condition	Total harmonic distortion in source current I_{THD} in (%)
Without SAPF	27.57
With conventional PI controller	6.65
With fractional PI controller	5.35

6 Conclusion

This paper shows the positive aspects of fractional PI controller over classical/conventional PI controller used with shunt active power filter to eliminate the harmonics and compensating reactive power of the system. The tremendous use of semiconductor (solid-state) devices causes various power quality issues because of their nonlinear characteristics. To overcome these power quality issues different methods have been used. In this paper, fractional PI controller has been studied and the performance of fractional PI controller has been compared with the performance of conventional PI controller. Results show that the performance of fractional PI controller is better than conventional PI controller as the harmonics reduces up to 6.65% by using CPI controller while the % of harmonic reduction is up to 5.35% by using FPI controller.

References

1. Singh, B., Al-Haddad, K., Chandra, A.: A review of active filters for power quality improvement. *IEEE Trans. Ind. Electron.* **46**(5) (1999)
2. Akagi, H.: Trends in active power line conditioners. *IEEE Trans. Power Electron.* **9**, 263–268 (1994)
3. Akagi, H., Fujita, H.: A new power line conditioner for harmonic compensation in power systems. *IEEE Trans. Power Deliv.* **10**, 1570–1575 (1995)
4. Peng, F.Z., Akagi, H., Nabae, A.: Compensation characteristics of the combined system of shunt passive and series active filters. *IEEE Trans. Ind. Appl.* **29**(1) (1993)
5. Fujita, H., Akagi, H.: A practical approach to harmonic compensation in power systems-series connection of passive and active filters. *IEEE Trans. Ind. Appl.* **21**(6) (1991)
6. Jain, S.K., Agarwal, P., Gupta, H.O.: A control algorithm for compensation of customer-generated harmonics and reactive power. *IEEE Trans. Power Deliv.* **19**(1) (2004)
7. Akagi, H., Nabae, A., Atoh, S.: Control strategy of active power filters using multiple voltage-source PWM converters. *IEEE Trans. Ind. Appl.* **IA-22**, 460–465 (1986)
8. Akagi, H., Atoh, S., Nabae, A.: Compensation characteristics of active power filter using multiseries voltage-source PWM converters. *Elect. Eng. Jpn.* **106**(5), 28–36 (1986)
9. Choe, G.-H., Park, M.-H.: A new injection method for ac harmonic elimination by active power filter. *IEEE Trans. Ind. Electron.* **35**(I) (1988)
10. Duke, R.M., Round, S.D.: The steady state performance of a controlled current active filter. *IEEE Trans. Power Electron.* **8**, 140–146 (1993)
11. Akagi, H., Kanazawa, Y., Nabae, A.: Instantaneous reactive power compensators comprising switching devices without energy storage components. *IEEE Trans. Ind. Applicat.* **IA-20**, 625–630, (1984)
12. Akagi, H., Nabae, A.: The p-q theory in three-phase systems under non sinusoidal conditions. *Eur. Trans. Elect. Power Eng.* **3**(1), 27–31 (1993)
13. Bhattacharya, A., Chakraborty, C., Bhattacharya, S.: Parallel-connected shunt hybrid active power filters operating at different switching frequencies for improved performance. *IEEE Trans. Ind. Electron.* **59**(11) (2012)
14. Gyugyi, L., Otto, R.A., Putman, T.H.: Principles, applications of static, thyristor controlled shunt compensators. *IEEE Trans. Power App. Syst* **97**, 1935–1945 (1978)
15. Gyugyi, L.: Reactive power generation, control by thyristor circuits. *IEEE Trans. Ind. Applicat.* **15**, 521–532 (1979)
16. Podlubny, I.: Fractional-order systems and PID controllers. *IEEE Trans. Autom. Control* **44**(1), 208–214 (1999)
17. Zou, Z., Zhou, K., Wang, Z., Cheng, M.: Fractional-order repetitivecontrol of programmable AC power sources. *IET Power Electron* **7**(2), 431–438 (2013)
18. Podlubny, I.: Fractional Differential Equations. Academic, San Diego, CA (1999)
19. Oldham, K.B., Spanier, J.: The Fractional Calculus. Academic, New York (1974)
20. Miller, K.S., Ross, B.: An Introduction to the Fractional Calculus and Fractional Differential Equations. Wiley, New York (1993)
21. Samko, S.G., Kilbas, A.A., Marichev, O.I.: Fractional Integrals and Derivatives and Some Of their Applications. Nauka i Technika, Minsk, U.S.S.R (1987)
22. Li, H., Luo, Y., Chen, Y.: A fractional order proportional and derivative (FOPD) motion controller: tuning rule and experiments. *IEEE Trans. Control Syst. Technol.* **18**(2) 2010
23. Jain, S.K., Agarwal, P., Gupta, H.O.: Design simulation and experimental investigations, on a shunt active power filter for harmonics, and reactive power compensation. *Electr. Power Compon. Syst.* **31**, 671–692 (2003) (Taylor and Francis)
24. Cois, O., Lanusse, P., Melchior, P., Dancla, F., Oustaloup, A.: Fractional systems toolbox for Matlab: applications in system identification and Crone CSD **2**, 7 (2014)

High Speed 64-Bit Booth Encoded Multiplier Using Compressor



G. M. G. Madhuri, Ch. Aruna Kumari, Ch. Monica
and N. Prasanthi Kumari

Abstract The present paper is about design methodology of High speed Booth Encoded Multiplier. A Booth Multiplier consists of the Encoder, the partial product tree, carry propagate adder. The multiplicand and multiplier size (n) is 64-bit unsigned operands. Radix-16 Booth recoded multiplier is implemented using VHDL. To lessen the partial product addition, compressors are used. Using 3:2, 4:2, 5:2, 6:2, 7:2 compressors, and carry save and propagate adder, all partial products are added to get the final output product. The multiplier is implemented in VHDL using Xilinx.

Keywords Radix-16 booth encoding multiplier · 7:2 compressors · VHDL

1 Introduction

In digital signal processing applications, the most vital components are Multipliers. Binary multiplication [1] consists of:

- (1) The multiplier is recoded in digits in a particular number system
- (2) Partial products are to be obtained by multiplication of each bit by the multiplicand
- (3) Using multi-operand addition methods reducing the partial product array and final carry summing of the two operands to acquire the final product result.

The general encoding practice translates a binary number into a signed operand with least redundant bits and assumes a vital job as it characterizes the quantity

G. M. G. Madhuri (✉) · Ch. A. Kumari
Department of ECE, PSCMR College of Engineering, Vijayawada, Andhra Pradesh, India
e-mail: gmgmadhuri@gmail.com

Ch. Monica
Research Scholar KL University, Vaddeswaram, India
e-mail: monicachatla@gmail.com

N. P. Kumari
University of Petroleum and Energy Studies, Dehradun, India

of the number of partial products [1, 2]. For radix- r the binary number consists of nonredundant radix- r bits and are encoded from $\{0, 1, \dots, r - 1\}$. For n-bit numbers, two's complement representation gives n/m partial products and $n + 1/m$ for unsigned representation.

Radix-four Booth multiplier [3] is frequently used as it encodes a binary number into radix-four signed digits using $\{-2, -1, 0, 1, 2\}$ and also requires simple shifts and complementation for producing partial products.

For production of the partial products higher radix signed recoding uses odd multiples of the multiplicand, i.e., obtained by not only simple shifts, but also need additions. In radix-16 encoding, the encoding set is $\{-8, -7, -6, -5, \dots, 0, \dots, 4, 5, 6, 7, 8\}$. Odd multiples of multiplicand i.e., $\times 3, \times 5$, and $\times 7$ are to be produced. X_3 is shifted to left by one bit to get $\times 6$. The production of the odd multiples needs a two-stage subtraction or addition.

2 Radix-16 Booth Multiplier

Let X be the multiplicand operand $X = \{x_{63} \dots \times 1, \times 0\}$ and Y be the multiplier $Y = \{y_{63} \dots y_1, y_0\}$. The preliminary way is the encoding of the multiplier operand [7], i.e., segregating them into four bit groups and encoded to bits in the set $\{-8, -7, \dots, 0, \dots, 7, 8\}$. For each term, 5 bits of the multiplier Y are taken and are encoded, i.e., 4 bits of 1 multiplier and 1 bit of the preceeding digit to determine the encoded digit. The encoder is designed using 16:1 MUX and 2:1 MUX to select one of the multiples of Y . The select lines are to select multiples from 0 to 15 and 2:1 to select sign bit.

The following table explains the encoding of bits (Table 1).

Figure 1 shows radix-16 booth encoder circuit. Three different size adders are to be used to get the odd multiples of Y . The following table shows the size of the adder. $6Y$ is the result of shifting operation of $3Y$, one position to the left. The multiples $Y, 2Y, 4Y$, and $8Y$ are obtained by shifting Y to the left by 0, 1, 2, or 3 bits, respectively (Table 2).

The different types of adders [7] used for addition of partial products are ripple-carry adders, carry-skip adders, look-ahead carry adders, and carry save adders. In this paper, partial products are summed up using compressors of different size.

3 Partial Product Tree Reduction

The partial product tree is reduced using compressors. The later stages are added using carry save adders, carry propagate adders. Half adders, 3:2 compressor, 4:2, 5:2, and 6:2, 7:2 [4–6] compressor are used.

3:2 compressor: In this, the inputs are i_1, i_2, i_3 and cin and outputs are sum and carry. This is designed using the following expressions:

Table 1 Encoding of bits

$X_{i+3}, X_{i+2}, X_{i+1}, X_i, X_{i-1}$	PP
00000	+X0
00001, 00010	+X1
00011, 00100	+X2
00101, 00110	+X3
00111, 01000	+X4
01001, 01010	+X5
01011, 01100	+X6
01101, 01110	+X7
01111	+X8
10000	-X8
10001, 10010	-X7
10011, 10100	-X6
10101, 10110	-X5
10111, 11000	-X4
11001, 11010	-X3
11011, 11100	-X2
11101, 11110	-X1
11111	X0

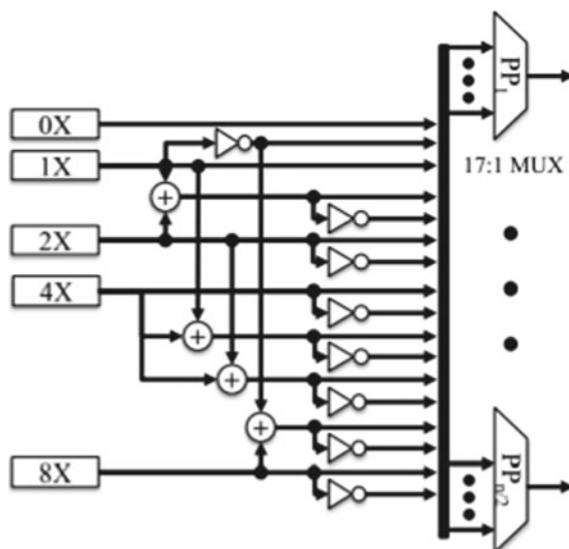
Fig. 1 Radix-16 booth encoder

Table 2 Partial products operation

PP	Operation
3Y	$Y + 2Y$
5Y	$Y + 4Y$
7Y	$8Y - Y$

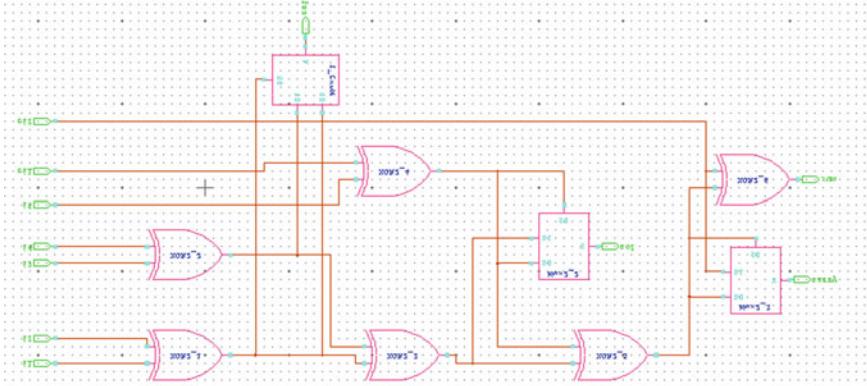


Fig. 2 5:2 Compressor circuit

$$\text{sum} = i1 \oplus i2 \oplus i3 \oplus \text{cin}$$

$$\text{carry} = (i1 * i2) + (i2 * \text{cin}) + (i2 * i3) + (i1 * \text{cin}) + (i1 * i3) + (i3 * \text{cin})$$

4:2 Compressor: In this, four partial products are compressed into two partial products. This block accepts four inputs i_1, i_2, i_3, i_4 and cin and produces three outputs sum, carry, and cout .

$$\text{sum} = i1 \oplus i2 \oplus i3 \oplus i4 \oplus \text{cin}$$

$$\text{cout} = (i1 \oplus i2)i3 + \overline{(i1 \oplus i2)}i1$$

$$\text{carry} = (i1 \oplus i2 \oplus i3 \oplus i4)\text{cin} + \overline{(i1 \oplus i2 \oplus i3 \oplus i4)}i4$$

5:2 Compressor: This block accepts five inputs, two carry inputs, and generates sum with two carry outputs. The circuit consists of ex-or gates and multiplexers. The circuit is shown in Fig. 2.

6:2 Compressor and 7:2 Compressor: This block accepts five inputs, two carry inputs and generates sum with two carry outputs. The circuit is shown in Fig. 3a and b.

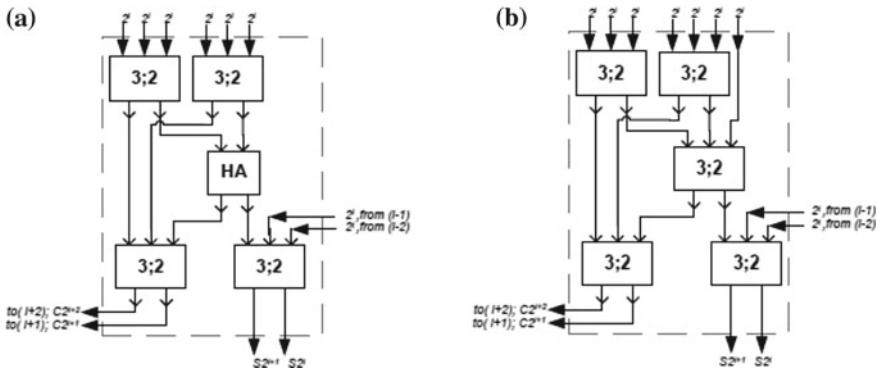


Fig. 3 a 6:2 Compressor b 7:2 Compressor

Stage -I of Addition														
P64	P65, P63	P6, P62					P6, P121	P5, P122	P4, P123	P3, P124	P2, P125	P1, P126	P0, P127	
7	7	7					7	6	5	4	3	2	1	7-7:2 compressors
7	7	7												6-6:2 compressors
2	1													5-5:2 compressors
Stage -II														
*	*						*	*	*	*	*			2-Half Adder
*	*						*	*	*	*	*			Carry save Adder
*	*						*	*	*					
*	*						*	*						
*	*													
Stage -III														
*	-	-					*	*	*					Carry propagate adder

Fig. 4 Partial product reduction

4 32×32 and 64×64 Multiplier

The 32×32 multiplier circuit is implemented using booth recoding technique. To generate partial products Radix-16 recoding is used. To reduce partial products array, different sizes of compressors are used. In the initial stage of reduction c_{in} is assumed to be zero. In later stage reduction is done by forwarding c_{in} for 32×32 multiplier sum of partial products is done in two stages. Three stages of partial product reduction are done in 64×64 multiplier to lessen the delay and increase the performance.

Partial product reduction for 64×64 multiplier is shown in Fig. 4.



Fig. 5 **a** Simulation result of 5:2 compressor **b** Simulation result of 64-bit multiplication

5 Conclusion

In this paper, a new Radix16 modified booth multiplier architecture is presented to execute the multiplication-accumulation operation, which is the key operation. The simulation results of 5:2 compressor and 64-bit multiplication are shown in Fig. 5. The slice LUTs for 32×32 multiplier are 2877 and max. Combinational delay is 45.64 ns. 128 IOs are used in this multiplier.

References

- Yeh, W.-C., Jen, C.-W.: High-speed Booth encoded parallel multiplier design. *IEEE Trans. Comput.* (2000)
- Kang, J.-Y., Gaudiot, J.-L.: A simple high- speed multiplier design. *IEEE Trans. Comput.* (2006)
- Kuang, S.-R., Wang, J.-P., Guo, C.-Y.: Modified Booth multipliers with a regular partial product array. *IEEE Trans. Circuit Syst.* (2009)
- Gopineedi, P.D., Thapliyal, H., Srinivas, M.B.: Novel and efficient 4: 2 and 5: 2 compressors with minimum number of transistors designed for low-power operations
- Bothra, R., Gulhane, P.: Power and delay analysis of 5–2 compressor. *Int. J. Adv. Technol. Innov. Res.* (2016)

6. Akoushideh, A., Najafi, A., Maybodi, B.M.: Modified architecture for 27:2 compressor. *Can. J. Electr. Electron. Eng.* **3**(8) (2012)
7. Saranya, A., Rajaram, M.: An optimization of complexity level in 32 bit vedic multiplier design using adder technique. *Int. J. Adv. Eng. Res. Dev. (IJAERD)*

Comparative Analysis of Least Square, Minimum Mean Square Error and KALMAN Estimator Using DWT (Discrete Wavelet Transform)-Based MIMO-OFDM System



Neha Awasthi and Sukesha Sharma

Abstract This paper includes the BER performance of least square, minimum mean square estimator and KALMAN filter is evaluated. Wireless channel experience the ill effect of multipath propagation, signal fading, tracking of channel, and distortion of the transmitted signal using channel estimation technique. DWT-based MIMO-OFDM system is implemented and channel estimation is done using LS, MMSE, and KALMAN filter. Channel estimation employed to recover the transmitted data. The method was tested by MATLAB simulation and completion was scrutinized with distinct modulations like BPSK (binary phase shift keying), QAM (quadrature amplitude modulation), PSK (phase shift keying) and QPSK (quadrature phase shift keying) of least square, KALMAN filter, and minimum mean square estimator.

Keywords Channel estimation · MIMO (multiple input multiple output) · OFDM (orthogonal frequency division multiplexing) · Pilot insertion · Least square · Minimum mean square estimator · KALMAN filter · BER · DWT · QAM · PSK · BPSK and QPSK

1 Introduction

Channel estimation has an imperative task to retrieve the transmitted data to ameliorate receiver performance [1]. A wavelet is a wave which has an amplitude that starts at zero, will ample and then decline back to zero. DW assists in acquiring apex rate transmission. In addition, fourier transform merely deals with the frequency component; moreover wavelets are functions that deal with time as well as in frequency domain [2]. Further, pilots inserted by using two approaches like comb and block type [1, 3]. In addition, comb kind pilot comprises an excessive degree of communica-

N. Awasthi (✉) · S. Sharma
UIET, Panjab University, Chandigarh, India
e-mail: neha.271awasthi@gmail.com

S. Sharma
e-mail: er_sukesha@yahoo.com

cation thus gives superior SNR [1, 4]. In this analysis, the signal is deteriorated into frequency elements and distinct modulation strategies are QAM, PSK, BPSK, and QPSK. The amount at which error prevail is termed as BER [2, 5]. To conclude, the BER performance of least square, minimum mean square estimator and KALMAN filter is evaluated. To conclude, the ramification furnished in the paper and simulated using MATLAB software.

2 Performance of the Discrete Wavelet Transform-Based MIMO-OFDM System

To commence, the system utilizes modulation and demodulation notion to assess the BER. In addition, the IDWT transfigures the statistics arrangement into the time estate waves further permute parallel into serial wave. On the other hand, tumult appended into communicated wave thus permutes the signal serial to parallel. In addition, for demodulation of the acquire signal the statistics is forwarded to the DWT segment. Further, DWT statistics recuperate and modify the signal into the frequency component. In addition, pilot signals are extricated; moreover the signal is demodulated and acquires binary data. The input statistics bits are amalgamated and depicted data symbols “ $\psi(n)$ ” are altered to design a signal “ $P(n)$ ” constituted as [1, 6]:

$$P(n) = \text{IDWT}\{\psi(n)\} \quad (1)$$

$$\psi(n) = e^{j2\pi n/T_u(t-T_s)} \quad (2)$$

where, “ $T(u)$ ” denoted sending time, “ t ” denoted time and “ T_s ” denoted overall time, “ $\psi(n)$ ” denoted mother wavelet. In other words, wave “ $P(n)$ ” can be constituted as [1, 6]:

$$P(n) = [P(0), P(1), \dots, P(N-1)] \quad (3)$$

where, “ N ” denoted subcarriers, “ $P(n)$ ” accounted frequency selective multipath fading channel constituted as [6]:

$$Z(n) = P(n) * b(n) + W(n) \quad (4)$$

where, “ $*$ ” denoted convolution, “ $b(n)$ ” denoted Impulse response, “ $W(n)$ ” denoted AWGN and “ $b(n)$ ” denoted channel impulse response constituted as [1, 6]:

$$b(n) = \sum_{v=0}^{v-1} b_v b_y(t - T_v) + W(n) \quad (5)$$

where, “ b_v ” denoted complex channel impulse response, “ v ” denoted propagation path, “ $W(n)$ ” denoted AWGN, “ T_v ” denoted propagation time, “ b ” denoted transmitted signal constituted as [6]:

$$b(n) = [b(0), b(1), \dots, b(N-1)] \quad (6)$$

Later serial to parallel transformation finished and constituted as [1, 6]:

$$Y(k) = DWT[P(n)] = (Y(n), \psi(n)) = 2^{-n/2} \sum_{n=\infty}^{\infty} (P(n)) \psi(2^{-n}n - k) \quad (7)$$

where, “ Ψ ” denoted mother wavelet and the signal “ $Y(k)$ ” comprised in frequency domain is constituted as [6]:

$$Y(k) = P(k)b(k) + W(k) \quad (8)$$

where, “ $P(k)$ ”, “ $W(k)$ ”, and “ $b(k)$ ” denoted DWT of “ $P(n)$ ”, “ $W(n)$ ”, and “ $b(n)$ ” and approximated channel response constituted as [6]:

$$P(k) = Z(k) / b(k) \quad \text{where, } k = 0, 1, \dots, (N-1) \quad (9)$$

3 Least Square Estimator, Minimum Mean Square Error Estimator, and KALMAN Filter

The LS is featured as the fraction between the input and the output data; moreover the LS has low intricacy [1]. By using a statistical function to establish a line of best fit by minimizing the sum of squares of errors and is taken to calculate channel impulse response of transmitted data is constituted as [1, 7]:

$$Ix = [I(0) I(1) I(2) \dots I(n-1)] \quad (10)$$

where, “ $I(n-1)$ ” shows “nth” bit from total transmitted bits and received signal constituted as [7]:

$$Jx = [J(0) J(1) J(2) \dots J(n-1)] \quad (11)$$

$$b(k) = Ix/Jx \quad (12)$$

MMSE estimator is considerably more complex than least square estimator while LS experiences high mean square error (MSE) [6, 8]. The minimum mean square estimator uses second-order statistics which involves using channel autocovariance in order to minimize the square error by using Eq. (4) (Fig. 1).

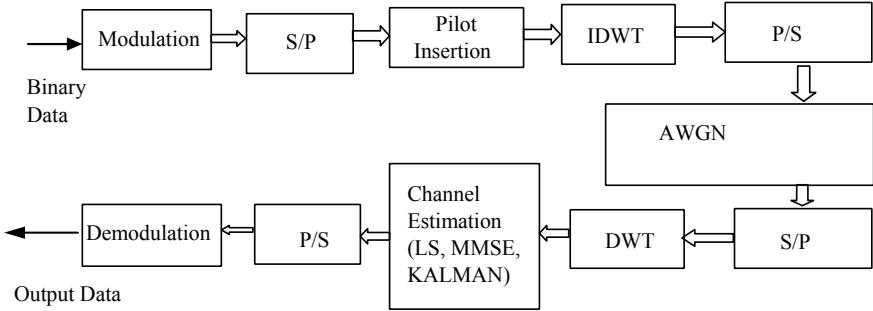


Fig. 1 Discrete wavelet transform based MIMO-OFDM system using LS, MMSE, and KALMAN [1]

$$\mathbf{y} = \mathbf{x} \mathbf{f} \mathbf{h} + \mathbf{W} \quad (13)$$

MMSE estimation stated as:

$$\mathbf{H}_{\text{MMSE}} = \mathbf{f} \mathbf{R}_{\mathbf{hy}} \mathbf{R} - \mathbf{I}_{\mathbf{yy}} \mathbf{y} \quad (14)$$

where,

$$\begin{aligned} \mathbf{R}_{\mathbf{hy}} &= E(\mathbf{hy}) = \mathbf{R}_{\mathbf{hh}} \mathbf{f}_{\mathbf{HxH}} \\ \mathbf{R}_{\mathbf{yx}} &= E(\mathbf{hy}) = \mathbf{x} \mathbf{f} \mathbf{R}_{\mathbf{hh}} \mathbf{f}_{\mathbf{HxH}} \end{aligned}$$

where, “ $\mathbf{R}_{\mathbf{yy}}$ ” denoted autocovariance matrix, “ $\mathbf{R}_{\mathbf{hy}}$ ” denoted the cross-covariance matrix and “ \mathbf{HxH} ” is the covariance matrix [6, 7]. KALMAN contains statistical calculations, moreover, to assess state estimate to diminish the MSE. In addition, to filter the impurities KALMAN filter is an efficacious approach [9]. Moreover, KALMAN utilized to preserve the estimations of past, present, and future condition. Prediction equation is called as predictor equation and time update equation [6, 9]. Update equation is called as measurement update equation and is also called as corrector equation and state space model equation is written as:

$$\mathbf{Y}_r = \mathbf{C}_r \mathbf{U}_r + \mathbf{W}_r(m) \quad (15)$$

$$\mathbf{U}_r = \mathbf{D}_r \mathbf{U}_{r-1} + \mathbf{E}_r \quad (16)$$

KALMAN gain is used to predict the system, to evaluate and control signals and is studied by predicted and measured states constituted as [1, 6]:

Predict states: $U(r) = CU(r - 1) + Du(r - 1)$

Predict error covariance: $S(r) = CS(r - 1)C^T + Q$

Compute the KALMAN gain: $K(r) = S(r)H^T (H^T S(r) H^T + R)^{-1}$

Estimate the states by correcting the predictions: $U(r) = U(r) + K(r)(z(r) - HU(r))$

Update error covariance: $S(r) = (I - K(r)H)S(r)$

4 Wavelet Transform (WT) and DWT-Based MIMO-OFDM

To commence, wavelets have the multiresolution capability in which DWT provides the examination of the signal in both times as well in the frequency domain [10]. Merely by substitution of fast fourier block with discrete wavelet transform [10]. The WT is additionally referred to as “mother wavelet” which is a prototype for the generation of other window functions [11]. In addition to there is no requirement to add cyclic prefix. Moreover, it furnishes the probe of the signal in distinct domain like time and frequency [12]. WT is located in both time and frequency domain as of acquires scaling and translation of a scaling function and wavelet function [13].

5 Ramification and Simulation

The comparison of the completion of LS, MMSE, and KALMAN using distinct modulations are illustrated in the statistics (Table 1).

As shown in Figs. 2,3 and 4 and Tables (2, 3, 4), for discrete wavelet-based MIMO-OFDM system the BER of KALMAN is least as compared to LS and MMSE (Fig. 5).

Table 1 Distinct channel frameworks for simulation

Framework	Value
Modulation	QAM, PSK, QPSK, BPSK
Subcarriers	64, 128, 256
No. of symbols	100
Constellation	64
Type of channel	AWGN/Rayleigh flat fading
SNR range	[0:2:40]

Fig. 2 Using QAM modulation BER of LS, MMSE, and KALMAN illustrated

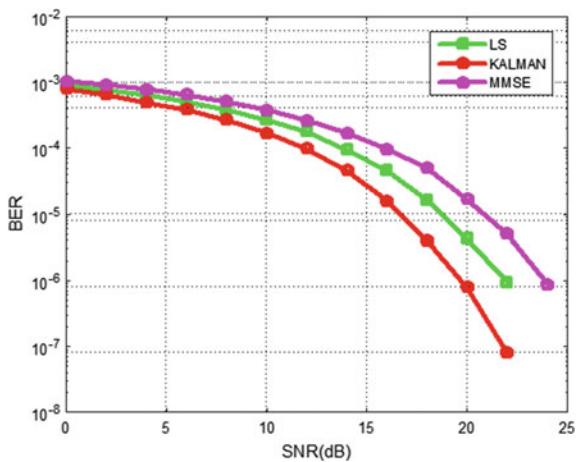


Fig. 3 Using PSK modulation BER of LS, MMSE, and KALMAN illustrated

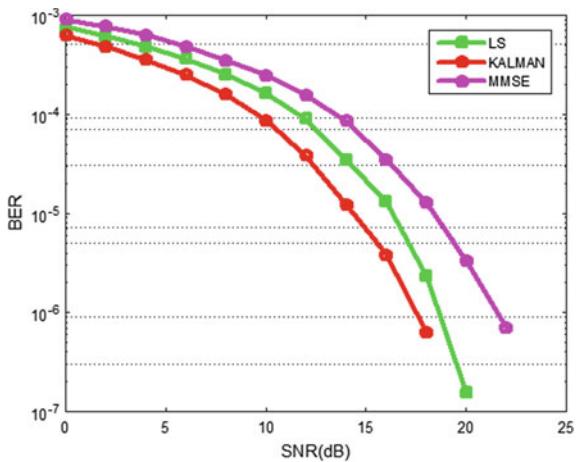


Fig. 4 Using BPSK modulation BER of LS, MMSE, and KALMAN illustrated

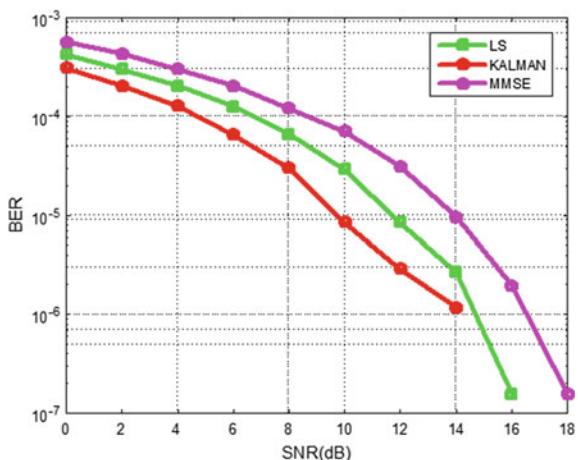


Table 2 In the Least Square the BER values of QAM, PSK, BPSK and QPSK values is shown

SNR (In dB)	Bit error rate			
	Least square			
	QAM	PSK	QPSK	BPSK
0	0.000929	0.000759	0.000584	0.000429
2	0.000774	0.000612	0.000450	0.000302
4	0.000649	0.000484	0.000329	0.000203
6	0.000515	0.000361	0.000216	0.000126
8	0.0005129	0.000252	0.000142	0.0000660
10	0.000267	0.000161	0.0000730	0.0000292
12	0.000179	0.0000900	0.0000360	0.00000843
14	0.0000937	0.0000345	0.0000113	0.00000265
16	0.0000461	0.0000129	0.00000304	0.000000156
18	0.0000160	0.00000234	0.000003046	0
20	0.00000429	0.000000156	0.000001640	0
22	0.00000937	0	0.000000390	0

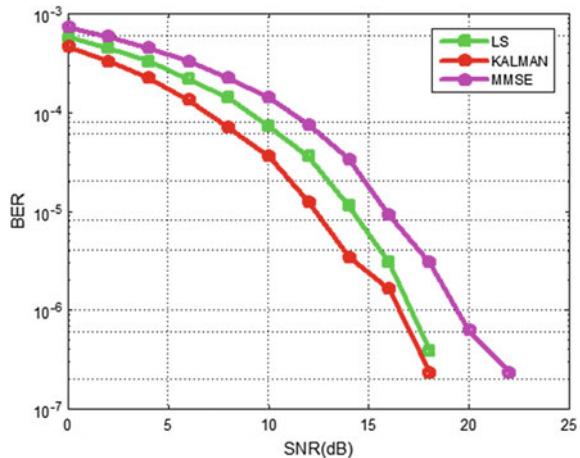
Table 3 In the KALMAN filter the BER values of QAM, PSK, BPSK and QPSK values is shown

SNR (In dB)	Bit error rate			
	KALMAN			
	QAM	PSK	QPSK	BPSK
0	0.00080828	0.000621	0.000459	0.000307
2	0.000651	0.000480	0.000328	0.000204
4	0.000493	0.000352	0.000223	0.000128
6	0.000387	0.000157	0.000134	0.0000655
8	0.000268	0.0000866	0.0000703	0.00000851
10	0.000170	0.0000380	0.0000363	0.00000289
12	0.0000978	0.0000121	0.0000124	0.00000117
14	0.0000460	0.00000382	0.00000343	0
16	0.0000156	0.000000625	0.00000164	0
18	0.00000390	0	0.000000234	0
20	0.000000781	0	0	0
22	0.000000782	0	0	0

Table 4 In the MMSE ESTIMATOR the BER values of QAM, PSK, BPSK and QPSK values is shown

SNR (In dB)	Bit error rate			
	MMSE			
	QAM	PSK	QPSK	BPSK
0	0.00104	0.000901	0.000729	0.000567
2	0.000925	0.000768	0.000585	0.000431
4	0.000786	0.000628	0.000449	0.000301
6	0.000652	0.000480	0.000328	0.000203
8	0.000512	0.000350	0.000223	0.000121
10	0.000382	0.000247	0.000142	0.0000702
12	0.000264	0.0000855	0.0000751	0.0000310
14	0.000170	0.0000347	0.0000339	0.00000953
16	0.0000959	0.0000125	0.00000914	0.00000195
18	0.0000496	0.00000335	0.00000304	0.000000156
20	0.0000168	0.000000703	0.000000625	0
22	0.00000500	0	0.000000234	0
24	0.000000859	0	0	0

Fig. 5 Using QPSK modulation BER of LS, MMSE, and KALMAN illustrated



6 Conclusion

To conclude, the communicated signal went through innumerable multipath fading, reflection, refraction, and propagation to the receiver side these consequences of channel need to diminish to retrieve the actual signal. To ameliorate the actual signal, effects of the channel needs to acquire to diminish at the receiver side. To ramifications, by employing discrete wavelet transform in MIMO-OFDM system the BER and the

throughput accomplishment is ameliorated. Thus, it perceived that by using QPSK, QAM, BPSK, and PSK modulations the achievement of KALMAN is superior and BER is least than LS and MMSE.

References

1. Awasthi, N., Sharma, S.: Comparative analysis of KALMAN and least square for DWT based MIMO-OFDM system. In: ICAICR. Springer (2018)
2. De Beek, V., Kate, S., Sandell, M., Wilson, S.K.: On channel estimation in OFDM systems, 0–5. <https://doi.org/10.1109/VETEC.1995.504981> (1995)
3. Asadi, A., Tazehkand, B.M.: A new method to channel estimation in OFDM systems based on wavelet transform. *Int. J. Digit. Inf. Wirel. Commun. (IJDIWC)* **3**(1), 1–9 The Society of Digital Information and Wireless Communications (2013). ISSN: 2225-658X
4. Zaier, A., Boualleque, R.: Channel estimation study for block type pilot insertion in OFDM system under slowly time-varying conditions **3**(6) (2011)
5. Kumar, L., Homayeun, N.: A review of wavelets for digital wireless communication. *Wirel. Pers. Commun.* **37**(3–4), 387–420 (2006) (Kluwer Academic Publishers)
6. Awasthi, N., Sharma, S.: Comparative analysis of least square and minimum mean square error estimator using DWT based MIMO-OFDM system. In: ICWSNUCA. Springer (2018)
7. Bouchlel, A., Sakle, A., Mansouri, N.: Performance comparison of DWT based MIMO-OFDM and FFT based MIMO-OFDM (2005)
8. Moqbel, M.A.M., Wangdong, Ali, A.Z.: MIMO channel estimation using the LS and MMSE algorithm. Hunan University, Changsha, Hunan, China. Hajjah University, Hajjah, Yemen IOSR. J. Electron. Commun. Eng. (IOSR-JECE) **12**(1), 13–22. e-ISSN: 2278-2834, ISSN: 2278-8735, Ver. II (2017)
9. Shrivats, A.K.: A comparative analysis of LS and MMSE channel estimation techniques for MIMO-OFDM system **1**(8), 44–48 (2015)
10. Panchal, L., Upadhyay, M., Shah, N., Amin, A.: Significant implementation of LS and MMSE channel estimation for OFDM technique **4**(3), 38–41 (2014)
11. Gupta, A.K., Pathela, M., Kumar, A.: Kalman filtering based channel estimation for MIMO-OFDM. *Int. J. Comput. Appl.* **53**(15), 0975–8887 (2012). Dehradun Institute of Technology, Dehradun
12. Anuradha, Kumar, N.: BER analysis of conventional and wavelet-based OFDM in LTE using different modulation techniques. In: Proceedings of Engineering and Computational Sciences (RAECS), UIET Panjab University Chandigarh, India, pp. 1–4. IEEE (2014)
13. Meenakshi, D., Prabha, S., Raajan, N.R.: Compare the performance analysis for FFT based MIMO-OFDM with DWT based MIMO-OFDM. In: ICECCN, pp. 441–445 (2013)

An ANP-GRA-Based Evaluation Model for Security Features of IoT Systems



Akshay Hinduja and Manju Pandey

Abstract The world has been experiencing a notable growth on the application of Internet of Things (IoT)-based appliances as IoT has been established as a major component of Information Technology after digital computer, Internet, and mobile technologies. However, the rapid growth of IoT-based systems also raises two undeniable concerns: security and privacy. Since IoT-based appliances are inherently exposed to the security attacks, security and privacy must be taken into account at the time of their acquisition. Owing to their economic concerns, the IoT-based system manufacturers may compromise with security issues, which can cause a breach of the network security and exploitation of user data. Therefore, an assessment framework for security features of IoT system is imperative. In this paper, we develop an assessment framework to evaluate the security features of IoT-based equipment. The proposed assessment framework uses hybrid multi-criteria decision making (MCDM) methodology. The framework comprises two widely used MCDM methods, Analytic Network Process (ANP) and Grey Relational Analysis (GRA). The evaluation process of security features of IoT system is twofold: first ANP is used to assign importance to the evaluating criteria, and then GRA is used to assess the alternatives with respect to the criteria. We also carried out an empirical study on assessment of IoT-based healthcare devices.

Keywords Internet of things (IoT) · Evaluation of IoT security features · Multi criteria decision-making (MCDM) · Analytic network process (ANP) · Grey relation analysis (GRA)

A. Hinduja (✉) · M. Pandey

Department of Computer Applications, National Institute of Technology Raipur, Raipur, Chhattisgarh, India

e-mail: hinduja.akshay@gmail.com

1 Introduction

Internet of Things (IoT) has become the fourth pillar of this era of Information Technology after digital computer, Internet, and mobile technologies. The notion of interconnecting “things” using sensors and communication means has become widespread now, and has influenced almost every industry and government organizations. The most prominent applications of IoT are *Smart Home, wearable healthcare devices, smart grid, smart farming, smart supply chain and smart retail* [1]. According to Cisco, the IoT technologies will be generating about \$9 trillion in annual sale connecting about 30.73 billion “things” to the Internet by 2020 [2]. This scenario unlocks great opportunities for researchers and entrepreneurs.

However, with great opportunities, the new technology also brings some gigantic challenges. In order to reap the benefits of IoT, service providers, developers, manufacturers, and end users must deal with these challenges. Two biggest concerns related to IoT are privacy and security. If we observe with technical point of view, high accessibility, omnipresent devices, and simplified network communication implies a system that is vulnerable to security attacks such as denial of services (DoS), spoofing, information disclosure and elevation of user privilege to accessibility [3, 4]. In addition, the omnipresence of wireless technologies for exchange of real-time information increases the possibility of unauthorised remote access, masquerading and eavesdropping.

Since IoT deals with sensitive information and has capability to influence not only virtual but physical environment also, the security of network and privacy of users must be regarded as the main concern while selecting IoT-based systems. In this paper, we build an evaluation model to assess performances of IoT-based systems over certain security features. First, we identify the major security requirements of the IoT system through literature survey. Afterwards, we integrate the Analytic Network Process (ANP) and Grey relational analysis (GRA) to evaluate the IoT equipment over security requirements. We leverage the ANP’s ability to establish relationships among criteria for accurate portrayal of inter-dependence and inner-dependence among criteria; we further use ANP to obtain criteria weights. GRA is used to evaluate IoT system with respect to security requirements/criteria.

The remainder of this paper is organized as follow: Sect. 2 provides a brief overview on preliminary concepts. Step by step methodology is explained in Sect. 3. In Sect. 4, we present an empirical study on evaluation of healthcare-related IoT systems. In the last section, we make concluding remarks.

2 IoT Architecture and Security

IoT can be regarded as a means to provide unique identifications to “things” so that they can communicate to one another and to human as well. Europe’s EPoSS organization defines the IoT as “Things having identities and virtual personalities

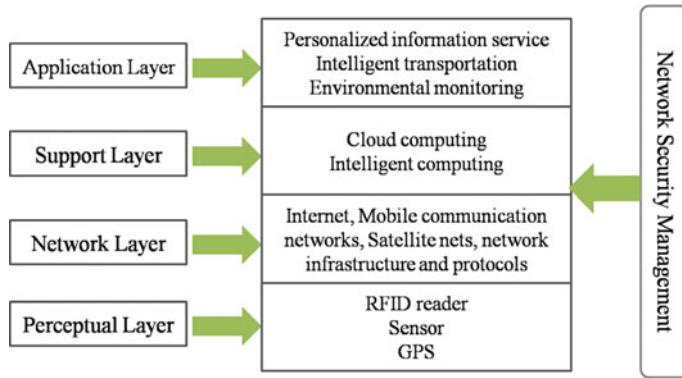


Fig. 1 Architecture of IoT [6]

operating in smart spaces using intelligent interfaces to connect and communicate within social, environmental, and user contexts.”

The IoT functions with 4-layer architecture, however, many researchers take 3-layer architecture into account [5]. The three core layers are: perceptual layer, which encompasses sensors and actuators, the network layer, which involves network-related devices such IPv6 enabled device, bluetooth and wi-fi emitters, and application layer, which contain application-specific programs and models. In addition, support layer is the most common adhered paradigm, which facilitates cloud computing and such additional services. Figure 1 shows a block diagram of 4-layer architecture of IoT.

Although the adoption of IoT by industries has been rapidly growing in recent years, it will be abandoned by end users if security and privacy are not ensured by service providers. In order to overcome the inherent susceptibility of IoT, lightweight security features are required to secure each layer of IoT. A single vulnerable layer may lead to unauthorized penetration to the overall IoT system. Some layer specific security requirements are shown in Fig. 2.

3 AHP-GRA-Based Evaluation Model

ANP method, developed by Saaty [7], is one of the most widely applied MCDM methods to managerial decision making. The recent applications of ANP are: ERP selection [8], CRM partner selection [9], selection of renewal energy [3], and evaluation of innovations [4]. Unlike AHP method which functions on certain unrealistic assumptions like independence of criteria from the other criteria, ANP enables the decision makers to establish interdependent and inner-dependent relationships among criteria with network relationship map (NRM). For detail procedure of ANP, please refer [7].

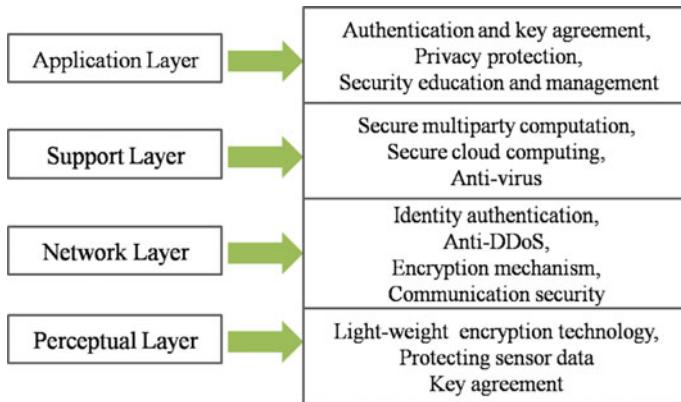


Fig. 2 Security requirements of different layers of IoT

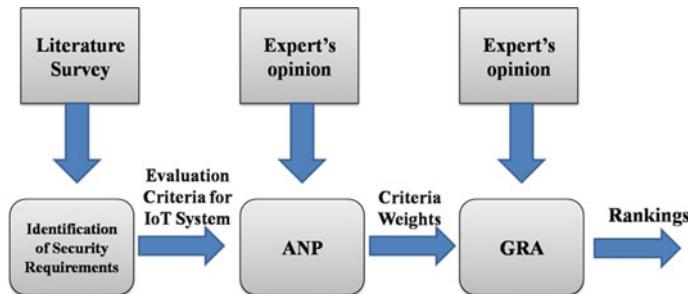


Fig. 3 Block diagram of evaluation model

GRA is a similarity-based MCDM method, which is derived from the grey theory [12]. It has been successfully applied to numerous industrial applications like selection of silicon wafer slicing machine [13], power distribution system [14] and evaluation of integrated-circuit marking process [15]. It measures the distances between the performances of alternatives over each criteria and the optimal performances; the less the distance the more the similarity. The alternative with the highest similarity with the optimal solution is considered to be the best available alternative. Please refer [16] for the comprehensive process of GRA.

In this section, we describe the evaluation model, which combines the merits of both MCDM methods, ANP, and GRA. The ANP is used to create Network Relationship Map (NRM) and obtain weights of security features while GRA is used to assess the IoT system alternatives over those security features. We provide a step by step methodology in Fig. 3.

Step 1: Identify decision elements. Define the goal, criteria, sub-criteria, and alternatives of the decision-making problem. The clusters of criteria, evaluating criteria and alternatives are obtained through extensive literature survey and experts' knowledge.

Step 2: *Construction of NRM.* Determine inter-dependence and inner-dependence among clusters using experts' opinion and extant literature, and create an NRM, which represents the decision problems and relationship among decision elements in the form of network. This NRM acts as a guide to the decision makers for further steps.

Step 3: *Pairwise comparisons and priority derivation.* After construction of NRM, obtain pair-wise comparisons among criteria within the clusters and criteria of inter-dependent clusters from decision-makers. Derive priorities of each pair-wise comparison matrix using Eigenvector method and check for their consistency. Matrices having more than 10% inconsistency should be sent to the decision makers for amendments.

Step 4: *Form the supermatrix.* Create a table and list all the evaluating criteria on it vertically and horizontally. Put in the weights derived from pair-wise comparisons of interdependent and inner-dependent clusters at the appropriate places of supermatrix.

Step 5: *Obtain the criteria weights.* After forming supermatrix, normalize it to obtain weighted supermatrix. Subsequently, raise the weighted supermatrix to a certain power so that the values in it become stable. The criteria weights can be obtained by normalizing the columns of the limit supermatrix.

Step 6: *Obtain the performance scores of alternatives.* Collect the performance score of alternatives through experts' opinion (for qualitative criterion) and brochures, and booklets (for quantitative criteria). Arrange the scores in a form of a matrix as each row of the matrix represents the performance scores of an alternative and each column represent performance scores of all alternatives over particular criterion. Subsequently, normalize the performance matrix using appropriate equation.

Step 7: *Calculate grey relational coefficient.* After normalizing the performance scores, define a reference sequence to measure similarity of alternatives with the hypothetical optimal alternative. Calculate the grey relational coefficients subsequently using Eq. (1). Determine the value of distinguishing coefficient (ρ) very carefully by consulting domain experts.

$$\xi(x_{0j}, x_{ij}) = \xi_{ij} = \frac{\delta_{min} + \rho\delta_{max}}{\delta_{ij} + \rho\delta_{max}}, \text{ for all } i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (1)$$

Step 8: *Selection of best alternative.* Compute the grey relation grade by synthesizing the grey relational coefficients of alternatives and corresponding criteria weights. Finally, select the alternative with the highest value of the grey relational grade.

4 Evaluation of IoT System for Security Features

In this section, we present an empirical study on assessment of security features of IoT systems. IoT has made an impact on every industry; no field of research and innovation is untouched from its influence. Owing to its limitless opportunities, healthcare is one of the most attractive application of IoT, however, security is a

crucial concern in health-related information [17]. Hence we choose IoT systems applied to healthcare domain for the evaluation of security features.

4.1 Identification of Decision Elements and Construction of NRM

Application of IoT to healthcare has been growing day by day. Primary applications are elderly care, fitness programs, remote health monitoring, and chronic diseases. But Security has always been an issue of IoT since the inception of its concept. For instance, What will happen if a IoT-based healthcare system faces Denial of Service (DoS) attack while collecting health-related information from critical patient ?? It can cost a life. Therefore, IoT-based healthcare systems must be invulnerable to any kind of security attacks and it is imperative to build an evaluation model for such systems.

In order to identify the security features, we make an extensive literature survey [6, 18–22]. We identify 15 security requirements of four fundamental layers (Table 1). We also validate the listed security requirements with a panel of experts, which consist of three researchers working on security of IoT. This expert panel provides their invaluable opinions throughout the work. Delphi method [23] is used to obtain consensus among experts.

After identifying the security requirements, we build an NRM with help of expert panel. In order to build the NRM, we analyze the inter-dependence and inner-dependence of security features distributed in four layers of IoT. We consider these layers to be clusters consisting of different security features or evaluating criteria for security of IoT systems. Figure 4 shows the NRM of evaluating criteria of security of IoT system applied to healthcare.

Table 1 Security features of different layers of IoT

Application layer (C1)	Support layer (C2)	Network layer (C3)	Perceptual layer (C4)
<i>C11:</i> Authentication & Key agreement	<i>C21:</i> User access control capability	<i>C31:</i> Anti-DDoS	<i>C41:</i> Intelligent node security
<i>C12:</i> Privacy protection	<i>C22:</i> Secure cloud computing	<i>C32:</i> Encryption mechanism	<i>C42:</i> Nodes information certification
<i>C13:</i> Platform security	<i>C23:</i> Information application security	<i>C33:</i> Communication security	<i>C43:</i> Anti-attack capability of WSN
	<i>C24:</i> Secure multiparty computation	<i>C34:</i> Application risk of IPv6	
		<i>C35:</i> Heterogeneous network recognition	

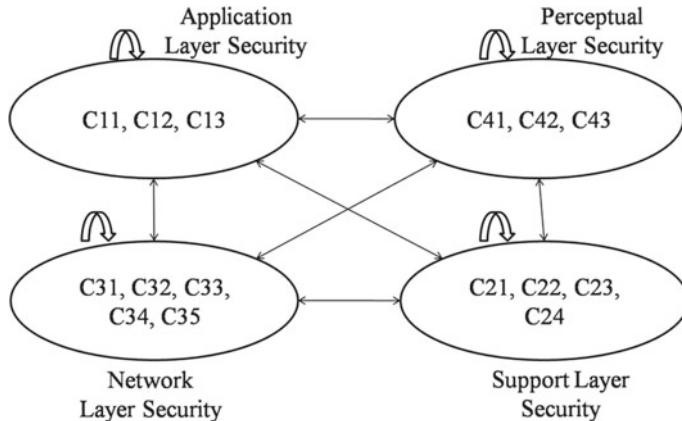


Fig. 4 NRM of evaluating security features

4.2 Deriving Criteria Weights Using ANP

The NRM shown in Fig. 4 provides the information about relationships among security features. The next step is to make pair-wise comparisons and derive priorities of security features with respect to other security features. Table 2 shows pair-wise comparison of security features of network layer with respect to authentication and key agreement (C11). Similarly, we obtain pair-wise comparisons of all criteria within their clusters with respect to all 15 criteria from expert panel and arrange their priorities to appropriate places of supermatrix (Table 3). Since sum of each column of supermatrix is generally more than one, therefore it is called un-weighted supermatrix. Afterwards, we raise the supermatrix to the significantly large power so that the values of the supermatrix become stable. Table 4 shows the limit supermatrix; in order to find weights of criteria, we need to normalize any column of the limit supermatrix.

Table 2 Pair-wise comparison of criteria of cluster C3 with respect to criterion C11

	C11	C31	C32	C33	C34	C35	Weights
C31	1	2	2	3	5	0.39	
C32	1/2	1	1	2	2	0.20	
C33	1/2	1	1	2	3	0.21	
C34	1/3	1/2	1/2	1	2	0.12	
C35	1/5	1/2	1/3	1/2	1	0.08	

Table 3 Un-weighted supermatrix

	C11	C12	C13	C21	C22	C23	C24	C31	C32	C33	C34	C35	C41	C42	C43
C11	0	1	0.40	0	0.50	0.25	0	1	0.25	0.33	0	0.75	1	0	0
C12	0.50	0	0.40	0	0.50	0.75	0	0	0.60	0.33	0	0.25	0	0	0
C13	0.50	0	0.200	0	0	0	0	0	0.15	0.33	0	0	0	0	0
C21	0.33	0	0	0.66	0.40	0.50	0.	0	0	0	0	0	0	0	0
C22	0.33	0	0	0	0.20	0.50	0.15	0	0	0	0	0	0	0.50	0
C23	0	0	0	0.33	0.20	0	0.60	0	0	0	0	0	0	0.50	0
C24	0.33	0	0	0	0.20	0	0.25	0	0	0	0	0	0	0	0.33
C31	0.39	0.75	0	0.15	0	0	0	0.20	0.20	0.20	0.20	0.20	0	0	0.66
C32	0.20	0	0	0	0	0	0	0.20	0.20	0.20	0.20	0.20	0.33	0	0
C33	0.21	0.25	0	0.25	0	0	0	0.20	0.20	0.20	0.20	0.20	0.33	0	0
C34	0.12	0	0	0.60	0	0	0	0.20	0.20	0.20	0.20	0.20	0.33	0	0
C35	0.08	0	0	0	0	0	0	0.20	0.20	0.20	0.20	0.20	0	0	0
C41	0	0.50	0	0	1	0.66	0.25	0	0	0	0	0	0.40	0.33	0.50
C42	0	0.50	0.33	0	0	0.33	0.60	0	0	0	0	0	0.40	0.33	0
C43	0	0	0.66	0	0	0	0.15	0	0	0	0	0	0.20	0.33	0.50

Table 4 Limit supermatrix

	C11	C12	C13	C21	C22	C23	C24	C31	C32	C33	C34	C35	C41	C42	C43
C11	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17
C12	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90
C13	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
C21	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.44	0.44
C22	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43	0.43
C23	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.34
C24	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32
C31	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11
C32	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
C33	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
C34	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
C35	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
C41	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
C42	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
C43	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04

4.3 Selection of Best Alternative Using GRA

In this section, we evaluate three IoT-based healthcare equipment on the basis of security features. For the sake of anonymity, we refer these healthcare solution alternatives as A, B, and C. First, we obtain the evaluation matrix from expert panel, which contains the performance scores of alternatives with respect to all the security

Table 5 Evaluation matrix

	C11	C12	C13	C21	C22	C23	C24	C31	C32	C33	C34	C35	C41	C42	C43
A	80	60	65	80	90	20	80	60	40	35	95	90	0	0	0
B	80	70	0	80	0	50	0	90	90	0	50	30	50	90	60
C	95	100	90	90	0	0	0	80	80	90	60	60	70	0	0

Table 6 Grey relational co-efficient matrix

	C11	C12	C13	C21	C22	C23	C24	C31	C32	C33	C34	C35	C41	C42	C43
A	0.71	0.56	0.59	0.71	0.83	0.39	0.71	0.56	0.46	0.44	0.91	0.83	0.33	0.33	0.33
B	0.71	0.63	0.33	0.71	0.33	0.50	0.33	0.83	0.83	0.33	0.50	0.42	0.50	0.83	0.56
C	0.91	1.00	0.83	0.83	0.33	0.33	0.33	0.71	0.71	0.83	0.56	0.56	0.63	0.33	0.33

Table 7 Grey relational grades of IoT systems

Healthcare IoT Solution	A	B	C
Final scores based on security features	0.613	0.526	0.695

features. Table 5 shows the evaluation matrix. Since all the criteria are qualitative, the experts provide their scores out of 100 and hence, normalization is not required. Since the evaluation scale is 0–100, it is also justified to choose a string of 100 s as a reference sequence. We compute the grey relational coefficients subsequently using Eq. (1), the grey relational coefficient matrix is shown in Table 6. The grey relation coefficients indicate the proximity of alternatives' performances to optimal (may be hypothetical) performances with respect to security features of IoT.

Afterward, we calculated the grey relational grade by synthesizing the grey relational coefficients with corresponding criteria weights and aggregate them. The final scores are provided in Table 7. The scores are not required to be normalized here; the greater the scores better the alternative. Alternative C in this empirical study obtains the highest grey relational grade; it implies that the alternative C is more reliable in terms of security features than other two alternatives.

5 Conclusion

The IoT has been one of the fastest growing technologies of the current era of information technology. The purpose of connecting “things” to “things” and human is to make things smart which in loop reinforce human capability. IoT has been enjoying the huge success, but it is also facing some challenges which incept from its inherent architecture, build to facilitate ease of access. Industries, trying to cease the opportunity of leveraging benefits of IoT may compromise with security enforcements. In order to ensure security of network and privacy of user data, the IoT systems must be evaluated for its security features.

This paper presents an evaluation model to assess the IoT system based on their performance with respect to security features. The paper also presents a case study on evaluation of IoT-based system for healthcare. The evaluation model uses ANP for obtaining weights of security features and GRA for selecting the most prominent alternative. The information used in MCDM methodology is gleaned from knowledge of domain experts.

For the future expansion of the study, fuzzy and advanced fuzzy methods can be utilized in order to obtain experts' opinions as linguistic variables. Furthermore, other prevalent MCDM methods as TOPSIS, VIKOR, and PROMETHEE can be used to evaluate different IoT devices. In addition, case studies can be carried out for other domains where IoT devices are used.

References

1. Gubbi, J., Buyya, R., Marusic, S., Palaniswami, M.: Internet of things (IoT): a vision, architectural elements, and future directions. *Futur. Gener. Comput. Syst.* **29**, 1645–1660 (2013)
2. Nedeltchev, P.: The internet of everything is the new economy. In: Cisco. https://www.cisco.com/c/en/us/solutions/collateral/enterprise/cisco-on-cisco/Cisco_IT_Trends_IoE_Is_the_New_Economy.html (2015)
3. Kidmose, E., Pedersen, J.M.: Security in internet of things. In: Cybersecurity and Privacy-Bridging the Gap, 99 p. River Publishers (2017)
4. Weber, R.H.: Internet of things-new security and privacy challenges. *Comput. Law Secur. Rev.* **26**, 23–30 (2010)
5. Sethi, P., Sarangi, S.R.: Internet of things: architectures, protocols, and applications. *J. Electr. Comput. Eng.* 2017 (2017). <https://doi.org/10.1155/2017/9324035>
6. Suo, H., Wan, J., Zou, C., Liu, J.: Security in the internet of things: a review. In: Proceeding of the 2012 International Conference on Computer Science and Electronics Engineering 2012, vol. 3, pp. 648–651 (2012). <https://doi.org/10.1109/ICCSEE.2012.373>
7. Saaty, T.L.: Decision Making with Dependence and Feedback: The Analytic Network Process : The Organization and Prioritization of Complexity, 2nd edn, Rws Publications (2001)
8. Selcuk, H., Zaim, S., Delen, D.: Selecting “The Best” ERP system for SMEs using a combination of ANP and PROMETHEE methods. *Expert Syst. Appl.* **42**, 2343–2352 (2015). <https://doi.org/10.1016/j.eswa.2014.10.034>
9. Büyüközkan, G., Güleryüz, S., Karpak, B.: A new combined IF-DEMATEL and IF-ANP approach for CRM partner evaluation. *Int. J. Prod. Econ.* **191**, 194–206 (2017). <https://doi.org/10.1016/j.ijpe.2017.05.012>
10. Büyüközkan, G., Güleryüz, S.: An integrated DEMATEL-ANP approach for renewable energy resources selection in Turkey. *Int. J. Prod. Econ.* **182**, 435–448 (2016). <https://doi.org/10.1016/j.ijpe.2016.09.015>
11. Chen, J.K., Chen, I.S.: Using a novel conjunctive MCDM approach based on DEMATEL, fuzzy ANP, and TOPSIS as an innovation support system for Taiwanese higher education. *Expert Syst. Appl.* **37**, 1981–1990 (2010). <https://doi.org/10.1016/j.eswa.2009.06.079>
12. Ju-Long, D.: Control problems of grey systems. *Syst. Control Lett.* **1**, 288–294 (1982). [http://dx.doi.org/10.1016/S0167-6911\(82\)80025-X](http://dx.doi.org/10.1016/S0167-6911(82)80025-X)
13. Lin, C.-T., Chang, C.-W., Chen, C.-B.: The worst ill-conditioned silicon wafer slicing machine detected by using grey relational analysis. *Int. J. Adv. Manuf. Technol.* **31**, 388–395 (2006). <https://doi.org/10.1007/s00170-006-0685-1>
14. Chen, W.-H., Tsai, M.-S., Kuo, H.-L.: Distribution system restoration using the hybrid fuzzy-grey method. *IEEE Trans. Power Syst.* **20**, 199–205 (2005). <https://doi.org/10.1109/TPWRS.2004.841234>

15. Jiang, B.C., Tasi, S.-L., Wang, C.-C.: Machine vision-based gray relational theory applied to IC marking inspection. *IEEE Trans. Semicond. Manuf.* **15**, 531–539 (2002). <https://doi.org/10.1109/TSM.2002.804906>
16. Kuo, Y., Yang, T., Huang, G.W.: The use of grey relational analysis in solving multiple attribute decision-making problems. *Comput. Ind. Eng.* **55**, 80–93 (2008). <https://doi.org/10.1016/j.cie.2007.12.002>
17. Islam, S.M.R., Kwak, D., Kabir, M.H., et al.: The internet of things for health care: a comprehensive survey. *IEEE Access* **3**, 678–708 (2015)
18. Mahmoud, R., Yousuf, T., Aloul, F., Zualkernan, I.: Internet of things (IoT) security: current status, challenges and prospective measures. In: 2015 10th International Conference for Internet Technology and Secured Transactions (ICITST) 2015 pp. 336–341 (2016). <https://doi.org/10.1109/ICITST.2015.7412116>
19. Granjal, J., Monteiro, E., Silva, J.S.: Security for the internet of things: a survey of existing protocols and open research issues **43**, 1–11 (2015)
20. Zhang, B., Zou, Z., Liu, M.: Evaluation on security system of internet of things based on Fuzzy-AHP method BT. In: 2nd International Conference on E-Business and E-Government, ICEE 2011, 6 May 2011–8 May 2011. E-bus E-Government (ICEE), pp. 2230–2234 (2011)
21. Park, K.C., Shin, D.H.: Security assessment framework for IoT service. *Telecommun. Syst.* **64**, 193–209 (2017). <https://doi.org/10.1007/s11235-016-0168-0>
22. Jing, Q., Vasilakos, A.V., Wan, J., et al.: Security of the internet of things: perspectives and challenges. *Wirel. Netw.* **20**, 2481–2501 (2014). <https://doi.org/10.1007/s11276-014-0761-7>
23. Okoli, C., Pawlowski, S.D.: The Delphi method as a research tool: an example, design considerations and applications. *Inf. Manag.* Elsevier **42**, 15–29 (2004)

A Comparative Study of 2D-to-3D Reconstruction Techniques



Surya Pandey

Abstract Image reconstruction is primarily an inverse problem, which directs the attention to reconstructing the original image from the given 2D image which may be blurred, consisting of some noise, and/or some damaged regions. The conversion to 3D involves finding out missing data. Rotation in conjunction with visualization from numerous perspectives of the image is facilitated by the third dimension for 3D reconstruction. Detailed and swift 3D reconstruction has established its applications in various fields such as Computer Vision, Medical Imaging, Virtual Reality, etc. The purpose of this paper is to bring attention to the thorough study of some state-of-the-art methods, which is used for fast 3D reconstruction. These methods include techniques, namely “Patch based Multi-View Stereo (PMVS), DAISY Descriptors, 3D point cloud, 3D Template matching, Recurrent Neural Network”. In the end, a comparative study of these methods based on various parameters is done to guide the readers to check the suitability of each approach.

Keywords 3D reconstruction · Patch-based multi-View stereo · DAISY descriptor · 3D point cloud · 3D template matching · Recurrent neural network

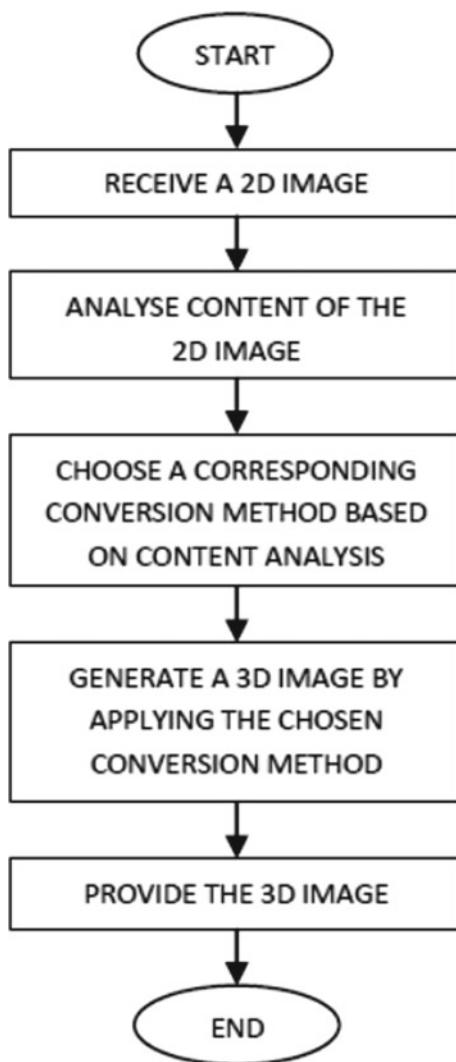
1 Introduction

Recent decade demands for 3D content for various purposes such as in movie industries, games, medical applications, etc. Numerous techniques have been accomplished, which are faster, efficient, and cost-effective. 3D reconstruction requires to be improved for a high enduring quality and performance. Various crux concerned with 3D reconstruction includes building reconstructed image from numerous 2D images. Face reconstruction methods require multiple images to be dealt with. The outcome leads to loss of memory and more time consumption. Shape from shading technique results in below par quality output on synthetic as well as real images.

S. Pandey (✉)

Orchids The International School, Sarjapur, Bangalore 562125, Karnataka, India
e-mail: suryaspandey@gmail.com

Fig. 1 Flowchart for 2D-to-3D conversion



Various earlier algorithms do not cope well with reconstructing images having poor texture or occlusion. A flowchart of 3D reconstruction is shown in Fig. 1.

3D reconstruction involves dealing with finding solutions to accuracy, occlusion, images involving sparse texture, etc. This paper presents solutions to these problems by surveying different works done by the authors. The paper is chalked out as Sect. 2 pays heed to the diligent analysis of each of the techniques mentioned. A comparative breakdown of these methods is done in Sect. 3. Section 4 brings an end to the paper.

2 Different Techniques of 3D Reconstruction

2.1 Patch-Based Multi-view Stereo

Reconstructing 3D models from multi-view images includes major applications such as in navigation, medical field, etc. 3D reconstruction from a confined number of stereo pairs captured by a 3D camera is an exacting task for Multi-view Stereo (MVS) reconstruction algorithms. A pioneering method of working built upon Patch-based Multi-view Stereo (PMVS) has been contemplated by Furukawa et al. [1]. The method targets at reconstructing 3D images using voluminous images. The considerable accomplishment of PMVS is that it is flexible. The given algorithm makes an effort to obtain a minimum of one patch in each image cell. To get a sparse set of patches first, the matching is done using Difference of Gaussian (DoG) and Harris operator, later, it is applied on regular pixels grid block of size $\beta \times \beta$ and expansion is performed. The sole objective of expansion is to reconstruct up to a certain extent, one patch in every image cell. Now, these set of patches are converted into surface meshes. Poisson surface reconstruction helps in mesh formation.

The major hindrance of this method was a terrific amount of time was spent in expansion step, which involved searching for more relevant patches. Wang et al. [2] devised a quicker way of patch expansion by manipulating quasi-dense matching [3] for finding dense initial patches using SIFT descriptor. This method is faster as much of the time is not wasted in the expansion step further reducing the computation time of the algorithm.

2.2 DAISY Descriptor

DAISY descriptor [4] gave rise to the evaluation of the dense depth maps for depth estimation accuracy, occlusion detection. Other feature descriptors were only accounted for local descriptors and did not handle the issue of occlusions. Due to large distortions and high occlusions involved, dealing with wide-baseline stereo has become a big challenge. One way to deal with this is to construct small patches. But the concern is it may stop to work for less textured images. SIFT and GLOH which are local region descriptors are designed to increase the robustness but it is only extended to small baseline stereo. Tola et al. [4] put forth DAISY descriptor that can be very well suited for wide-baseline stereo. This is done by making use of Expectation–Maximization(EM) [5]. The basic difference between SIFT, GLOH, and DAISY is that 3D histograms are computed again and again for each feature point in SIFT and GLOH, whereas the histograms are computed just once per region for DAISY descriptor. The histograms are then reused again for the neighboring pixels. This results in the faster computation of descriptors with an involvement of small overhead.

The EM strategy used was not accurate enough to reconstruct occluded images. Xeu et al. [6] introduced photometric discrepancy function emanate from DAISY descriptor. This method tries to reduce these problems to a certain extent by initiating a sparse set of patches by DAISY descriptor dependent on photometric discrepancy function (PDF).

The photometric discrepancy function for DAISY descriptor cannot accurately identify reliable patches having specular highlights or obstacles. To deal with this, DAISY adopts a method wherein it assumes that if the value of PDF is smaller than a given threshold, the pair of pixels is pitch-perfect. Shen [7] made use of just two images to build a reliable 3D image making the computation speed much faster than DAISY and PMVS.

2.3 3D Point Cloud

A point cloud includes a bunch of points which are used for 3D reconstruction. A method for 3D reconstruction based on point cloud studied here is space carving [8]. Space carving method feels the necessity for a calibrated digital camera along with rotating platform.

Volume intersection deduces the object shape. The method actually separates the object from the background by figuring out the outer contour line for every viewport. By combining all the shape lines obtained from every angle, 3D point cloud can be estimated. The space carving method takes a sequence of images from 360° for image and can construct the 3D model with the help of a sequence of outer contour lines obtained. The image is rotated for certain angle θ and the contour line is obtained for each rotated image. After obtaining the 3D points, redundant points are eliminated by removing the points that lie outside the cone created from the 3D mesh. This method is facile in reconstructing the object. At the same time, it is also cost-effective, but cannot be used for the large scene. The outer contour line obtained of the 2D image is translated into a 3D cone with the formula given in [9]. This method is good for reconstructing the object.

2.4 3D Template Matching

Although a lot has been emphasized on reconstructing 3D images, reconstruction of images which have poor texture and occlusion has not been looked upon so well. For partially occluded and/or minimal texture surfaces, the problem becomes even more severe as there is hardly any useful piece of information available about major parts of it [10] focuses on 3D shape recovery when a reference image and a single corresponding 3D template shape are known beforehand. A template image T in addition to the rest of the shape of the distort surface are given as inputs.

Occlusion is managed by making use of a relevancy score. This score is estimated a priori, which is later used to weigh the pixels when handling occlusion. A continuous-valued score is computed, which predicts the importance of pixels based on their relevancy score. The relevancy score is measured as the greatest value of NCC over a series of patches surrounding image location. Image regions containing little or no texture have low relevancy scores. Though the method has been devised for managing occlusion and sparse texture, they must be initialized or require supervised learning.

2.5 Recurrent Neural Network

An uncharacteristic RNN architecture has paved its way in [11] called as 3D Recurrent Reconstruction Neural Network (3D-R2N2). The network assesses out a mapping from objects' images to their respective 3D shapes from a wide collection of synthetic data using Deep CNN. The network receives as input, at least one or more object images from arbitrary viewpoints and outputs a reconstructed object as 3D occupancy grid.

It is impacted by the accomplishment of Long Short-Term Memory (LSTM) [12] networks apart from recent progress in single-view 3D reconstruction using CNN [13, 14]. The network has three components—(1) 2D Convolutional Neural Network, (2) 3D Convolutional LSTM, and (3) 3D Deconvolutional Neural Network.

The 2D-CNN encodes each input image, thus bringing it down to low-dimensional features. It is followed by a series of 3D-LSTM units, which either selectively renew their cell states or hold on to the states. In the end, 3D-DCNN scrutinizes the hidden states of the LSTM units. This helps in acquiring 3D probabilistic voxel reconstruction. The perk of this method is that the network trains and reconstructs without knowing the object category. This method is thought to be the best up till date. It goes beyond the modern methods for single-view reconstruction enabling 3D reconstruction of objects. This approach wants a slight supervision in learning process.

3 Comparative Analysis

After exploring the aforementioned techniques, a comparative analysis has been done based on various parameters listed below. It has been analyzed well to come to a conclusion, which method is suited best based on the parameters mentioned. Table 1 provides an overall summary of these techniques. A comparative scrutiny has been done with respect to parameters such as:

1. Number of images;
2. Occlusion;
3. Memory consumption;
4. Texture;

Table 1 Comparative analysis of 3D image construction techniques

Methods→ Parameters ↓	PMVS	DAISY descriptor	3D point cloud	3D template	RNN
No. of images	15–300	10–100	9	2	1
Occlusion	Not suitable	Suitable	Suitable	Suitable	Suitable
Memory consumption	Large memory	Large memory	Large memory	Less memory	Less memory
Texture	Not suitable	Better than PMVS	Suitable	Suitable	Suitable
Robustness	Less robust	Robust than PMVS	Robust	Robust	Robust
Accuracy	Accurate than DAISY descriptor	Accurate	Accurate	Accurate	Accurate

5. Robustness;
6. Accuracy.

3.1 Number of Images

1. PMVS: 15–300 number of images;
2. DAISY DESCRIPTOR: 10–100 number of images;
3. 3D POINT CLOUD: 9 images used [9];
4. 3D TEMPLATE MATCHING: 2 images (1 template and 1 input image);
5. RECURRENT NEURAL NETWORK: 1 image to maximum any number of images.

3.2 Occlusion

1. PMVS: It does not deal well with occlusion [15]. Occluded regions in the image are not visible enough in a number of images. So, these regions are eliminated because they may not fulfill the constraint of photometric discrepancy function. Also, it considers parallelly multiple current matches and propagates in a larger area neglecting the occluded regions.
2. DAISY DESCRIPTOR: Makes use of EM algorithm.
3. 3D POINT CLOUD [16]: 3D point cloud can manage occlusion very well. Pixels are always visible in an order that captures all occlusion relations between the entire set of pixels. Each iteration guarantees that if a pixel p occludes another

- pixel q when viewed from a camera, p will necessarily be visited before. Hence, gives a better result.
4. **3D TEMPLATE MATCHING:** Occlusion can be very well dealt with 3D template matching. Relevancy score is computed checking for occlusion. The score is high for un-occluded regions because it is consistent with the template. Score is low for occluded regions since it is not consistent with the template.
 5. **RECURRENT NEURAL NETWORK:** The network retains the states of the other parts but reforms the LSTM states for the precedently occluded sections.

3.3 Memory Consumption

1. **PMVS:** More memory consumption. It aims to use all the images available simultaneously. Hence, suffers from scalability issue when more number of images are used [7].
2. **DAISY DESCRIPTOR:** Less memory consumption [7]. This is obvious because it loads only two images for computation at a time.
3. **3D POINT CLOUD:** Memory consumption is large due to the fact that more number of images are used. The camera is rotated to get the images of an object from every possible angle storing more number of images.
4. **3D TEMPLATE MATCHING:** Less memory consumption since only two are images used, i.e., template and input image.
5. **RECURRENT NEURAL NETWORK:** Requires memory consumption as it gets trained from the saved dataset.

3.4 Texture

1. **PMVS:** Cannot deal with fine texture [17]. Patch-to-patch propagation is used which may overlook occluded regions. This is hardly workable for matching different scenes.
2. **DAISY DESCRIPTOR [17]:** Suited for textured regions as it makes use of EM algorithm.
3. **3D POINT CLOUD:** Suited well for textured regions.
4. **3D TEMPLATE MATCHING:** Good for textured regions because of the use of relevancy score that is computed. Relevancy score is small for highly textured regions, which means the template does not match with the given image sequence.
5. **RECURRENT NEURAL NETWORK:** Independent of accurate feature correspondences or adjacent camera viewpoints as required by other methods. It works best for texture.

3.5 Robustness

1. PMVS: Not robust than DAISY DESCRIPTOR [7];
2. DAISY DESCRIPTOR: Faster than PMVS [7];
3. 3D POINT CLOUD: Not robust;
4. 3D TEMPLATE MATCHING: Robust;
5. RECURRENT NEURAL NETWORK [11]: The network is robust enough to produce the reconstructed image.

3.6 Accuracy

1. PMVS [7]: More accurate than DAISY;
2. DAISY [7]: Less Accurate;
3. 3D POINT CLOUD: Accurate;
4. 3D TEMPLATE MATCHING: Accurate;
5. RECURRENT NEURAL NETWORK: Accurate enough from the remaining methods discussed.

Gazing at Table 1, we can deduce that as far as the number of images is concerned, all except PMVS and DAISY descriptor perform very well and so does memory consumption. Occlusion is dealt perfectly by 3D point coud, template matching, and RNN. All except PMVS does a great job for 3D reconstruction associated with texture. The same goes with accuracy and robustness.

4 Conclusion

The paper exhibits a detailed study of various state-of-the-art techniques that are able to construct 3D images effectively and efficiently. Also, a comparison between the methods is made based on various parameters to analyze its scope. Different methods mentioned here have their own added advantages. It will show beyond doubt to be of one's best interest for many researchers in future. It can be put to good use in future for widening its applications which may be limited now.

References

1. Furukawa, Yasutaka, Ponce, Jean: Accurate, dense, and robust multiview stereopsis. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(8), 1362–1376 (2010)
2. Wang, L., Chen, R., Kong, D.: An improved patch based multi-view stereo (PMVS) algorithm. In: 3rd International Conference on Computer Science and Service System. Atlantis Press (2014)

3. Ito, K., Xiong, K.: Gaussian filters for nonlinear filtering problems. *IEEE Trans. Autom. Control.* **45**(5), 910–927 (2000)
4. Tola, E., Lepetit, V., Fua, P.: An efficient dense descriptor applied to wide-baseline stereo. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(5), 815–830 (2010)
5. Strecha, C., Fransens, R., Van Gool, L.: Combined depth and outlier estimation in multi-view stereo. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (2006)
6. Xue, B., et al.: A DAISY descriptor based multi-view stereo method for large-scale scenes. *J. Vis. Commun. Image Represent.* **35**, 15–24 (2016)
7. Shen, S.: Accurate multiple view 3D reconstruction using patch-based stereo for large-scale scenes. *IEEE Trans. Image Process.* **22**(5), 1901–1914 (2013)
8. Kutulakos, K.N., Seitz, S.M.: A theory of shape by space carving. *Int. J. Comput. Vis.* **38**(3), 199–218 (2000)
9. Shi, X., Zhao, B.: An algorithm for 3D reconstruction based on point-cloud image sequence. In: 2015 IEEE Advanced Information Technology, Electronic and Automation Control Conference (IAEAC) (2015)
10. Haines, B., Bai, L.: Towards a robust patch based multi-view stereo technique for textureless and occluded 3D reconstruction. In: Computational Vision and Medical Image Processing V: Proceedings of the 5th Eccomas Thematic Conference on Computational Vision and Medical Image Processing (VipIMAGE 2015, Tenerife, Spain, 19–21 Oct 2015). CRC Press (2015)
11. Choy, C.B. et al.: 3d-r2n2: a unified approach for single and multi-view 3d object reconstruction. In: European Conference on Computer Vision. Springer International Publishing (2016)
12. Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural Comput.* 1735–1780 (1997)
13. Eigen, D., Puhrsch, C., Fergus, R.: Depth map prediction from a single image using a multi-scale deep network. *Advances in Neural Information Processing Systems* (2014)
14. Liu, F., Shen, C., Lin, G.: Deep convolutional neural fields for depth estimation from a single image. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2015)
15. Haines, B., Bai L.: Towards a robust patch based multi-view stereo technique for textureless and occluded 3D reconstruction. In: Computational Vision and Medical Image Processing V: Proceedings of the 5th Eccomas Thematic Conference on Computational Vision and Medical Image Processing (VipIMAGE 2015, Tenerife, Spain, 19–21 Oct 2015). CRC Press (2015)
16. Hammoudi, K. et al.: Generating occlusion-free textures for virtual 3D model of urban facades by fusing image and laser street data. In: 2012 IEEE Virtual Reality Workshops (VRW). IEEE (2012)
17. Fischer, J. et al.: A feature descriptor for texture-less object representation using 2D and 3D cues from RGB-D data. In: 2013 IEEE International Conference on Robotics and Automation (ICRA). IEEE (2013)

Recent Research Trends and Methods Used in Moving Object Detection



Tannistha Pal, Arindam Deb, Anwesha Roy and Praseed Debbarma

Abstract Object detection is one of the main challenges in the field of image processing from the beginning. Many methods had been established and proposed through the years, but none of them could give a high accuracy rate along with low computation time. Background subtraction is a key technique in object detection in security and surveillance purpose with a high accuracy rate, but it comes in a cost of high computational time. Frame differencing came with low computational time, but it cannot give us the accuracy that we want. Optical flow and temporal differencing are independent of irregularities of light features, but accuracy is not up to the mark. The main obstacles of detecting an object in a video are still present in proposed techniques till date. In this paper, we mainly describe the recent trends and works done in the field of object detection. We have discussed various important methods of object detection and pointed out their positive aspects and limitations.

Keywords Object detection · Video surveillance · Object classification

1 Introduction

Accurate detection of moving object is a fundamental task for computer vision-based applications such as intelligent transportation systems, airport security systems, video monitoring systems, and so on, where background subtraction is widely

T. Pal (✉) · A. Deb · A. Roy · P. Debbarma

Department of Computer Science and Engineering, National Institute of Technology,
Agartala, Barjala, Jirania 799055, India

e-mail: tannisthapaul@gmail.com

A. Deb

e-mail: arindamdeb2013@gmail.com

A. Roy

e-mail: pratibhalaya96@gmail.com

P. Debbarma

e-mail: pappu09.pd@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

S. Choudhury et al. (eds.), *Intelligent Communication, Control and Devices*,
Advances in Intelligent Systems and Computing 989,
https://doi.org/10.1007/978-981-13-8618-3_28

adopted [1–4]. The trending applications of object detection methods are mostly to satisfy the purpose of automatic surveillance. Demands of the traditional, as well as conventional research topics of computer vision, are obvious. They are providing enormous features in the field of moving object detection for security (guard for significant buildings), pattern recognition (spotting endangered species), and many more. One of the simplest and popular techniques for moving object detection is a background subtraction method, which uses background modeling concept, but its computational time is more in detecting objects [1]. Frame differencing is also a popular technique for detecting objects but it won't be adequate to give high accuracy results in detecting the objects. Temporal difference method is very simple and it can detect objects in real time, but it fails to detect the object against illumination change. Optical method also gives excellent result in object detection but its computation cost is very noise and it is sensitive to noise. The main focus of the paper is to provide a little contribution toward the future implementation of previously published research papers by providing a brief idea on their respective contents. Section 2 comprises of analytical review on various approaches of object detection and their techniques. We have also done a comparative analysis of the recent research trends and put all the shortcomings and positive aspects of the papers as well as our findings. In Sect. 3, we have concluded the paper along with the future works in the field of moving object detection incorporating various traditional methods.

2 Traditional Methods Used in Object Detection

Object detection is one of the most researched areas in the field of image processing and computer vision. Through the years, many works have been done, and many works are still running in this field [5–7]. The main objective is to accurately detect a moving object from a video. The researchers face a lot of problems during computing object detection algorithms. Their main challenges are sudden illumination change, dynamic background, camera moving, the shadow of the original object coming into the output frame, and many more. Generally, the object detection algorithms can be specified into four major methods. They are Frame Differencing, Background Subtraction, Optical Flow, and Temporal Differencing. None of the methods are good in both the field of accuracy and computational time. They lack in at least one of them. We have discussed the methods briefly in the below paragraph.

2.1 *Frame Differencing*

Kartika and Mohamed [8] used the Frame Differencing technique on consecutive frames along with post-processing Adaptive Threshold technique and Shadow Detection technique. Three techniques combined to give an adequate output where the object in motion is detected with a distinct shape and highly accurate dimension.

In the technique used in the paper [8], all the three techniques are interlinked with each other, and all three techniques should be performed in a sequential manner to obtain the desired result. Frame differencing won't be adequate to give high accuracy results. The work done in paper [9] used the same frame differencing technique, but instead of taking consecutive frames, here, three frames have been taken, one of which is taken as background and other two frames consist of the object in motion. Here, two parallel frame differencing is done. One, with the first background frame and the second frame with an object in motion and second, with the second frame with an object in motion and the third consecutive frame with an object in motion. Then, the two different frames are subjected to logical "OR" operation, and if the "OR" option gives intensity values 0, it is background or else, it is considered as foreground. This technique does not give a distinct dimensional figure in the output rather provides us with an estimated figure.

2.2 *Background Subtraction*

This is a well-known technique in the field of object detection. The main advantage of this method is that it gives us high accuracy rates than the other methods. For this reason, it is used to analyze footage found by remote surveillance cameras. But this high accuracy comes with high computational time. For this reason, it cannot be used in the real-time processing of a video. Dynamic background change can cause a lot of problems in object detection by this method. An object detection framework has been proposed in the paper [1] for surveillance and navigation applications, which combines many crucial steps. First, a background pixel model, which can be initialized from a single frame. Second, a pixel classification process, which classifies a pixel according to its corresponding model. Third, an update mechanism, which ignores the insertion time of a pixel in the pixel model and replaces values randomly.

2.3 *Optical Flow*

In this method, the complete sense of the optical flow field of the image is required. It gives excellent result but in the cost of a large amount of calculation of the optical flow field. It is also very sensitive to noise. Clustering is done in this method based on optical flow information obtained from the image.

2.4 *Temporal Differencing*

Murali and Girisha [10] have found the basic working of temporal differencing, which comprises of three consecutive frames selection and comparison of their varying

pixel intensities. It has been shown in the paper [11] that temporal differencing proves to be beneficial in various ways of being independent of the irregularities of light and weather conditions. It is invariant of color characteristics chosen by the user. It is highly recommended in case of motorized camera tracking. In the paper of Foreground Segmentation using GMM combined Temporal Differencing [12], the authors have used GMM for foreground and background subtraction and Temporal Differencing for marking the boundary of the target object.

A comparative analysis of recent research trends in detecting an object is summarized in Table 1.

3 Conclusion and Discussion

Object detection in a video is an efficient research field. It has a lot of application in today's world especially in the field of computer vision applications. Foreign object detection in surveillance camera is one of them. The researchers in this field are continuously working to develop an algorithm, which can efficiently detect the moving objects and will take low computational time. But they face a lot of problems in doing so. Camera movement, sudden illumination change, and shadow of a moving object create a lot of problems in detecting the real moving object. The output images contain a large amount of noises present. A moving object present in the initialization frame may become responsible for the creation of ghost. For these reasons, post-processing of output images is very much needed for later applications. In this field background subtraction, it always attracts researchers to work more on that method as it promises to give highly accurate results, but it lacks in computational speed which is the main reason it is not used in real-time scenarios. Though some of the background subtraction techniques takes lesser amount of time. Also, ghost suppression is very fast in this method. Other approaches though take less computational time, but accuracy rate is not up to the mark (e.g., frame differencing). Optical flow mainly takes a lot of time to calculate a huge amount of optical flow information of the image found from the video frame. For this reason, it has low computational speed. Temporal differencing though independent of irregularities of light features but is has low accuracy rate in many cases and it cannot remove the shadow of the moving object. So far, we have come into conclusion that any method lonely is not enough for our need of high accuracy rate and low computational time.

So, we hope that in the near future, the abovementioned traditional methods may incorporate with each other and will take the best of each method and will develop a new hybrid method, which will help to meet our needs.

Table 1 A comparative analysis of recent research trends

Technique used	Author/year	Advantages	Shortcomings	Computational time	Dataset used	Remark
Background subtraction	Hofmann et al [13]	1. It performs best for the categories like the baseline, shadow and thermal 2. Median filtering technique used reduced noise	1. Poor results may be generated if general lighting conditions in the scene change with time	Slow	Not provided	This technique basically helps in human-computer interaction. But due to its slow computational speed, it is very hard to use it in real-time applications
	Xiang et al. [14]	It can deal with illumination change and distinguish shadows of the moving object	If it cannot handle sudden illumination change	Slow	Not provided	The main goal of the technique was to detect the vehicle properly
	Hao et al. [15]	It has high computational speed and can deal with slow background change	If the feature of background is like that of the object, then it cannot work properly	Fast	Not provided	It uses edge ratio that allows to detect object and shadow separately
	Maddalena and Petrosino [16]	1. Compatible with background changing 2. It is vigorous against false detection	Cannot be useful in case of sudden illumination change	Slow	BMC 2012 Benchmark	It is the enhanced version of traditional three-frame differential method
	Barnich and Van Droogenbroeck [1]	1. Fast ghost suppression 2. Intrinsic withstand to camera shake 3. Noise recovery technique applicable	1. Its computation speed is not as fast as others 2. It is not robust to large camera movements and lightning changes	Slow	PETS and HOUSE	The accuracy rate of this technique is far better than many other background subtraction techniques, but the main problem is its slow processing time
	Zivkovic [17]	1. Modeling of the background as a statistical (random) phenomenon rather than constant helps easier and faster classification of background pixel 2. Provides fast recovery	1. Satisfactory and suitable results won't be achieved if the background is bimodal 2. Scenes containing many, slowly moving objects cannot be exactly detected	Slow	Not provided	It is widely used in traffic monitoring, i.e., for counting the vehicle, detecting, and tracking vehicles
	Manzanera and Richefeu [18]	1. Faster computation 2. Robust and accurate detection of moving objects for a small cost in memory consumption	Adaption capability to certain complex scenes like a sudden change in brightness is poor. In case of very low motion, this method is likely going to take that as a background characteristic	Fast	Not provided	This technique is used to detect persons using the wall-mounted cameras. As its computational time is very less, so it is also used for many real-time applications

(continued)

Table 1 (continued)

Technique used	Author/year	Advantages	Shortcomings	Computational time	Dataset used	Remark
	Choi et al. [19]	1. Eradicate incorrectly detected foreground pixels caused by illumination 2. Brightness change of the pixels traced using only four sample pixels	1. Cannot remove shadows 2. Cannot overcome camera shaking		i-LIDS parked vehicle videos	The use of chromaticity difference model and a brightness ratio model made the proposed technique produce an excellent result
Frame differencing	Kartika and Mohamed [8]	By using the post-processing technique, the accuracy of the output increases and the shadow detection technique helps to merge the shadow with the background	The accuracy of frame differencing technique alone is quite low	Medium	Not provided	This is one most simplified and an efficient technique for object detection with a quite high accuracy rate. It provides adequate output without much distortion or shadow for object detection in an open environment
	Srivastava et al. [9]	Logical “OR” operation performed on the differenced image combines the object movement information in either past or future frame and frequently used for gathering an averaging effect of a multitude image of the similar scene to drastically reduce additive noise	This method does not provide an adequate dimension of the moving object, and the exact motion of the object cannot be traced properly	Fast	Not provided	This technique has given much the more accurate result than preexisting two-frame differencing techniques and background subtraction method, but lack in providing us with the exact description of the moving object in terms of Dimension and structure
	Chaohui et al. [5]	1. High recognition rate 2. Unique use of edge detection technique gives an improved result	Cannot overcome complicated background problems	Very fast	Not provided	In the proposed technique, the Canny detector is used to detect edges of two continuous frames, which are then divided into small blocks for deciding if they are moving objects by comparing the number of nonzero pixels
Temporal differencing	Triwan et al. [12]	1. It has high adaptability with dynamic scene changes 2. Noise effect is mostly erased using a median filter	1. Constraints of distinguishing common pixels in background and object 2. It detects only the background of the object	Slow	Standard datasets publicly available over the internet	The proposed technique amalgamates the outcomes of temporal differencing and that of GMM to accomplish eradicating the inappropriate regions
	Murali and Girisha [10]	It manages to detect the moving object even if the foreground and background are of similar color	1. Poor functioning in separating out all the actual feature pixels 2. Cast shadow area increases if regions not linked accurately during spatial clustering	Fast	PEITS video of 2001, 2004, 2006	This method has successfully provided a solution to the orifice created by the temporal differencing outcomes with the combination of simple spatial cluster and image fusion

(continued)

Table 1 (continued)

Technique used	Author/year	Advantages	Shortcomings	Computational time	Dataset used	Remark
Liang et al. [11]		1. The method is simple 2. The effect in outcomes due to ranging illuminations levels is very minute	1. The moving target is not detected as a solid structure 2. This method fails in detecting the object if there are no changes in three consecutive frames	Rapid	Experimented on many image sequences were taken during the day in complex weather conditions (sunny, rainy, snowy, windy days)	This process has majorly outperformed in their outcomes during complex weather conditions
Wang et al. [20]		Background model initialization is not required and easily handle camera jitter	Cannot distinguish between actual moving object and their shadow	Fast	Not provided	The background model is updated by revising the best matching weight vector
Optical flow	Lin et al. [21]	1. The accuracy rate of this technique is very high around 85% 2. It can deal with the main obstacles of object tracking	A poor performance against deformation, Illumination change, and out of view	Slow	Not provided	One of the major disadvantages is that it works mostly on assumption that the object's illumination change is not obvious

References

1. Barnich, O., Van Droogenbroeck, M.: Vibe: a universal background subtraction algorithm for video sequences. *IEEE Trans. Image Process.* **20**(6), 1709–1724 (2011)
2. Alawi, M.A., Khalifa, O.O., Islam, M.R.: Performance comparison of background estimation algorithms for detecting moving vehicle. *World Appl. Sci. J.* **21**(Mathematical Applications in Engineering), 109–114 (2013). ISSN 1818–4952, IDOSI Publications, 2013, <https://doi.org/10.5829/idosi.wasj.2013.21.mae.99934>
3. Nascimento, J.C., Marques, J.S.: Performance evaluation of object detection algorithms for video surveillance. *IEEE Trans. Multimed.* **8**(4) (2006)
4. Goyette, N., Jodoin, P.M., Porikli, F., Konrad, J., Ishwar, P.: Changecoloration.net: a new change detection benchmark dataset. In: 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 16–21, June 2012
5. Zhan, C., Duan, X., Xu, S., Song, Z., & Luo, M.: An improved moving object detection algorithm based on frame difference and edge detection. *IEEE* (2007)
6. Elhabian, S.Y., El-Sayed, K.M., Ahmed, S.H.: Moving object detection in spatial domain using background removal techniques—State-of-art. *Recent. Pat. S Comput. Sci.* **1**, 32–54 (2008)
7. Del-Blanco, C.R., Jaureguizar, F., García, N.: An efficient multiple object detection and tracking framework for automatic counting and video surveillance applications. *IEEE Trans. Consum. Electron.* **58**(3), 857–862 (2012)
8. Kartika, I., Mohamed, S.S.: Frame differencing with post-processing techniques for moving object detection in outdoor environment. *IEEE* (2011)
9. Srivastav, N., Agrwal, S.L., Gupta, S.K., Srivastava, S.R., Chacko, B., Sharma, H.: Hybrid object detection using improved three frame differencing and background subtraction. In: 7th International Conference on Cloud Computing, Data Science & Engineering—Confluence (2017)
10. Murali, S., Girisha, R.: Segmentation of motion objects from surveillance video sequences using temporal differencing combined with multiple correlation. *Adv. Video Signal Based Surveill. J.* (2009)
11. Liang, K., Wang, J., Zhao, T.: Variation temporal differencing for moving target detecting and tracking. In: International Conference on Electronics, Communications and Control (ICECC) (2011)
12. Tiwari, V., Chaudhary, D., Tiwari, V.: Foreground segmentation using GMM combined temporal differencing. In: International Conference on Computer, Communications, and Electronics (Comptelix). Malaviya National Institute of Technology Jaipur, 01–02 July 2017
13. Hofmann, M., Tiefenbacher, P., Rigoll, G.: Background segmentation with feedback: the pixel-based adaptive segmenter. In: Proceedings of IEEE Work-shop on Change Detection (2012)
14. Xiang, J., Fan, H., Liao, H., Xu, J., Sun, W., Yu, S.: Moving object detection and shadow removing under chang-ing illumination condition, pp. 1–10. Hindawi Publishing Corporation, Mathematical Problems in Engineering, Feb 2014
15. Hao, J., Li, C., Kim, Z., Xiong, Z.: Spatiotemporal traffic scene modeling for object motion detection. *IEEE, Intell. Trans. Syst.* (2012)
16. Maddalena, L., Petrosino, A.: The 3dSOBS + algorithm for moving object detection. *Comput. Vis. Image Underst.* **65**–73 (2014)
17. Zivkovic, Z.: Improved adaptive gaussian mixture model for back-ground subtraction. In: Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004, vol. 2 (2004)
18. Manzanera, A., Richefeu, J.C.: A new motion detection algorithm based on Σ -background estimation. *Pattern Recognit. Lett.* **28**, 320–328 (2007)
19. Choi, J.M., Chang, H.J., Yoo, Y.J., Choi, J.Y.: Robust Moving Object Detection Against Fast Illumination Change. Elsevier (2011)

20. Wang, Z., Liao, K., Xiong, J., Zhang, Q.: Moving object detection based on temporal information. *IEEE Signal Process. Lett.* **21**(11), 1404–1407 (2014)
21. Lin, L., Liu, B., Xiao, Y.: An object tracking method based on CNN and optical flow. In: 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD). IEEE (2017)

Simple Way of Tuning Centralized PI Controller for an Interacting Pilot Plant Distillation Column



R. Janani and I. Thirunavukkarasu

Abstract A simple robust control technique based on centralized PI is proposed for an interacting pilot plant distillation column. The diagonal elements of the overall open loop transfer function of the system are reduced to first-order plus dead-time (FOPDT) model. The proposed method considers Steady-State Gain Matrix and loop transfer function parameters like time delay and time constant for the centralized controller design. The proposed controller method is validated using Wood and Berry distillation column and pilot plant binary distillation column. The servo response shows that the designed centralized PI controller has good robust stability and performance.

Keywords Centralized control · MIMO process · Steady-state gain · PI control

1 Introduction

All process industries have many manipulated and controlled variables. For an MIMO systems, the loop interaction and existence of dead time and time delay creates a challenging task in designing a controller. Because of this interacting behavior, the MIMO process can be controlled by multi-loop controller or centralized controller. In this current research, a simple method of designing a centralized PI controller is proposed for an interacting pilot plant distillation column [1].

A typical distillation column used in the chemical industries is to separate a mixture of components into two or more products based on their volatility. In the

R. Janani

Department of Electronics and Instrumentation Engineering, Sri Chandrasekharendra Sarawathi Viswa Maha Vidyalaya, SCSVMV, Enathur, Kanchipuram 631561, TamilNadu, India
e-mail: [janurang@gmail.com](mailto:jануранг@gmail.com)

I. Thirunavukkarasu (✉)

Department of Instrumentation and Control Engineering , Manipal Academy of Higher Education, Manipal 576104, Karnataka, India
e-mail: it.arasu@manipal.edu

present research, a mixture of Isopropyl alcohol and water are considered for the distillation. The manipulated variables are the reflux flow rate (L) and reboiler power rate (Q), whereas the controlled variables are the tray temperatures T5 and T1 [2, 3].

This paper is systematized as follows: Sect. 2 describes the design methodology for Proportional-Integral Controller. Section 3 represents the design and simulation results of the proposed PI controller for both Wood Berry model and Lab-Scale Pilot Plant Distillation column model followed by conclusion.

2 Design of PI Controller

Consider an MIMO processes given as

$$G(s) = \begin{bmatrix} g_{11}(s)e^{-\tau_{11}s} & g_{12}(s)e^{-\tau_{12}s} \\ g_{21}(s)e^{-\tau_{21}s} & g_{22}(s)e^{-\tau_{22}s} \end{bmatrix} \quad (2.1)$$

The overall open-loop transfer function is

$$h_{11}(s) = g_{11} - g_{12}g_{22}^{-1}g_{21} \quad (2.2)$$

$$h_{22}(s) = g_{22} - g_{21}g_{11}^{-1}g_{12} \quad (2.3)$$

These overall open-loop transfer function obtained are approached into an FOPDT model given in Eq. (2.4) and its parameters are obtained using Eqs. (2.5)–(2.7).

$$h_{ii}(s) = \frac{K_{ii}e^{-L_{ii}s}}{T_{ii}s + 1} \quad (2.4)$$

$$K_{ii} = h_{ii}(0) \quad (2.5)$$

$$T_{ii} = \sqrt{\frac{K_{ii}^2 - |h_{ii}(j\omega_{cii})|^2}{|h_{ii}(j\omega_{cii})|^2\omega_{cii}^2}} \quad (2.6)$$

$$L_{ii} = \frac{\pi + \tan^{-1}(-\omega_{cii}T_{ii})}{\omega_{cii}T_{ii}} \quad (2.7)$$

where K_{ii} is the open loop gain, T_{ii} and L_{ii} are the time delay and time constant, respectively [1].

The centralized PI controller matrix $G_c(s)$ is given by

$$G_c(s) = \begin{bmatrix} g_{c11}(s) & g_{c12}(s) \\ g_{c21}(s) & g_{c22}(s) \end{bmatrix} \quad (2.8)$$

The controller $G_c(s)$ is calculated using an inverse steady-state gain matrix [2, 4] and the open-loop transfer function parameters as given below

$$G_c(s) = \left(K_C + \frac{K_I}{s} \right) * [G(s = 0)]^{-1} \quad (2.9)$$

where $G(s = 0)$ is the steady-state gain matrix, the proportional gain (K_C) and integral gain (K_I) are given by

$$K_C = \delta_1 \left(\frac{T_{smallest}}{L_{Largest}} \right) \quad (2.10)$$

$$\tau_I = \delta_2 L_{Largest} \text{ and } K_I = \frac{K_C}{\tau_I} \quad (2.11)$$

The largest delay and smallest time constant is considered here for the PI controller design. δ_1 and δ_2 are the tuning parameters with respect to proportional gain and integral time constant, respectively.

3 Simulation Results and Analysis

3.1 Example 1: Wood and Berry Model

The WB Column model is extensively used in literature for the study of different control techniques. The model transfer function is given by [5, 6]

$$G(s) = \begin{bmatrix} \frac{12.8}{16.7s+1} e^{-s} & \frac{-18.9}{21s+1} e^{-3s} \\ \frac{6.6}{10.9s+1} e^{-7s} & \frac{-19.4}{14.4s+1} e^{-3s} \end{bmatrix} \quad (3.1)$$

The steady-state gain matrix and its inverse are obtained as

$$G(s = 0) = \begin{bmatrix} 12.8 & -18.9 \\ 6.6 & -19.4 \end{bmatrix}$$

$$G(s = 0)^{-1} = \begin{bmatrix} 0.1570 & -0.1529 \\ 0.0534 & -0.1036 \end{bmatrix} \quad (3.2)$$

The open-loop transfer function of the WB model after considering the loop interaction is obtained in the form of FOPDT model as [4]

For Loop 1

$$q_1 = \frac{6.37e^{-1.36s}}{5.19s + 1} \quad (3.3)$$

For Loop 2

$$q_2 = \frac{-9.65e^{-3.49s}}{4.25s + 1} \quad (3.4)$$

The PI controllers setting is based on the largest time delay and smallest time constant of the loop transfer function. Accordingly, $L_{largest} = 3.49$ and $T_{smallest} = 4.25$. Assume $\delta_1 = 0.5$ and $\delta_2 = 1.5$. Thus, the PI tuning parameters are obtained as $K_C = 0.6089$ and $K_I = 0.1163$.

From Eq. (2.9), the controller matrix is obtained as

$$G_c(s) = \left(0.6089 + \frac{0.1163}{s} \right) \times [G(s = 0)]^{-1} \quad (3.5)$$

The centralized PI controller is given by

$$G_c(s) = \begin{bmatrix} \frac{0.0955s+0.0182}{s} & \frac{-0.0931s-0.0177}{s} \\ \frac{0.0322s+0.00616}{s} & \frac{-0.063s-0.012}{s} \end{bmatrix} \quad (3.6)$$

The servo response is evaluated by a step change in input r_1 and r_2 . Figure 1 shows the response of Y_1 interaction response Y_2 for a unit step change in r_1 and the response of Y_2 interaction response Y_1 for a step change in r_2 . This response shows that the controller provides proper tracking with the given setpoint. The Table 1 depicts the performance index for the servo response for WB model by the proposed method.

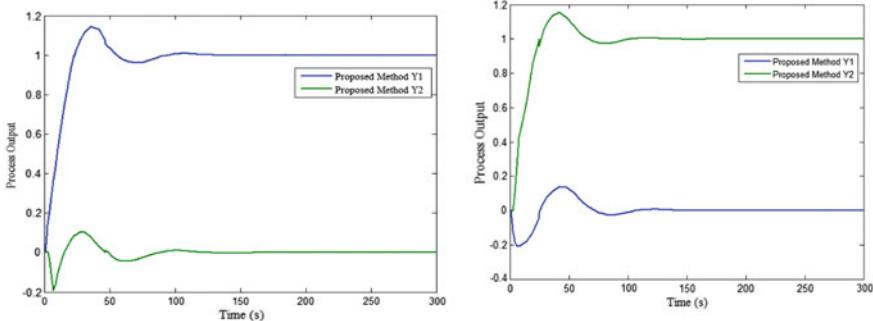


Fig. 1 Servo response of WB model for setpoint $r_1 = 1$ and $r_2 = 0$, setpoint $r_1 = 0$ and $r_2 = 1$ of proposed method

3.2 Example 2: Lab-Scale Pilot Plant Distillation Column

The real-time experimental setup of the lab-scale pilot plant distillation column is fabricated as shown in Fig. 2. The open-loop test is performed on the column and the empirical FOPTD model is considered for simulation studies identified by Vinaya and Arasu [2]. The process model is

$$G(s) = \begin{bmatrix} \frac{-0.13}{1.14s+1} e^{-0.03s} & \frac{0.18}{0.64s+1} e^{-0.03s} \\ \frac{-0.34}{1.23s+1} e^{-1.22s} & \frac{0.18}{0.32s+1} e^{-0.03s} \end{bmatrix} \quad (3.7)$$

Note that the time constants and dead times in the pilot plant distillation column model are measured in terms of hours and the process gain unit is given by °C/>. Since for the sampling time of 0.01 s is used in the VDPID-03 DAQ card, the open-loop data has huge data collection. Hence, the data were converted into hours to get the plant model (Table 2).

Table 1 Performance index for servo response by proposed method

Time domain characteristics	Y_{11}	Y_{21}	Y_{12}	Y_{22}	Main action	Interaction	Overall
IAE	14.02	4.615	8.001	15.78	29.8	12.616	42.416
ISE	7.063	0.3607	0.9932	8.214	15.277	1.3539	16.6309
ITAE	260.3	177.2	287	308.8	569.1	464.2	1033.3
ISTE	45.83	8.824	25.41	58.37	104.2	34.234	138.434

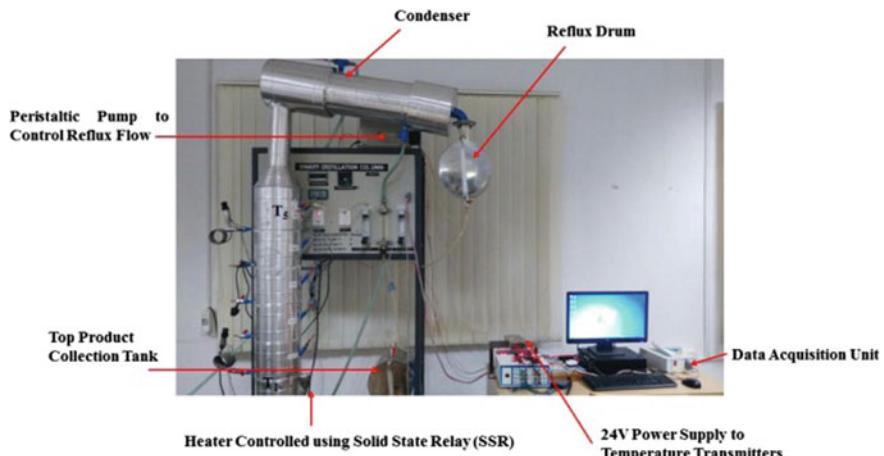


Fig. 2 Lab-scale pilot plant distillation column setup

Table 2 Performance index for servo response by proposed method

Time domain characteristics	\mathbf{Y}_{11}	\mathbf{Y}_{21}	\mathbf{Y}_{12}	\mathbf{Y}_{22}	Main action	Interaction	Overall
IAE	2.438	0.972	0.8681	9.489	11.927	1.8401	13.7671
ISE	8.268	0.729	0.0589	7.97	16.238	0.7879	17.0259
ITAE	12.17	8.51	7.72	72.63	84.8	16.23	101.03
ISTE	15.06	3.203	0.2618	25.57	40.63	3.4648	44.0948

The inverse steady-state gain matrix is obtained as

$$G(s = 0)^{-1} = \begin{bmatrix} 4.762 & -4.762 \\ 8.994 & -3.439 \end{bmatrix} \quad (3.8)$$

The open-loop transfer function of the VA model after considering the loop interaction is obtained in the form of FOPDT model as

For Loop 1

$$q_1 = \frac{0.21e^{-2s}}{1.35s + 1} \quad (3.9)$$

For Loop 2

$$q_2 = \frac{-0.29e^{-1.545s}}{0.735s + 1} \quad (3.10)$$

The PI controllers setting is based on the largest time delay and smallest time constant of the loop transfer function. Accordingly, $L_{largest} = 2$ and $T_{smallest} = 0.735$. Assume $\delta_1 = 1.5$ and $\delta_2 = 0.6$.

Thus, the PI tuning parameters are obtained as $K_C = 0.5513$ and $K_I = 0.4594$. From Eq. (2.9), the controller matrix is obtained as

$$G_c(s) = \left(0.5513 + \frac{0.4594}{s} \right) \times [G(s = 0)]^{-1} \quad (3.11)$$

The centralized PI controller is given by

$$G_c(s) = \begin{bmatrix} \frac{2.62s+2.18}{s} & \frac{-2.62s-2.18}{s} \\ \frac{4.956s+4.128}{s} & \frac{-1.895s-1.578}{s} \end{bmatrix} \quad (3.12)$$

The Table 2 depicts the performance index for the servo response for VA model by the proposed method.

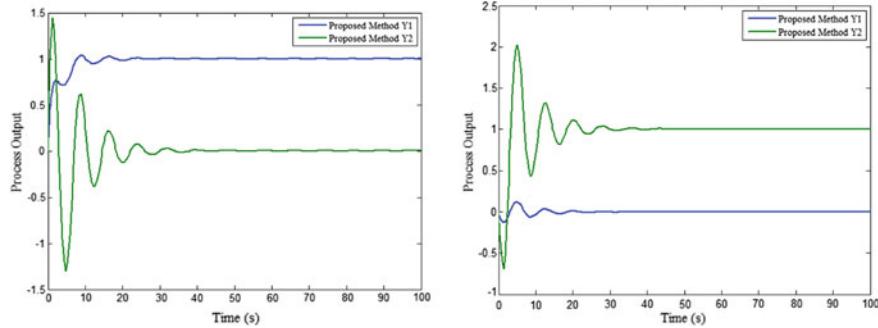


Fig. 3 Servo response of VA model for setpoint $r_1 = 1$ and $r_2 = 0$, setpoint $r_1 = 0$ and $r_2 = 1$ of proposed method

4 Conclusion

The simulation result shows that the closed-loop performance is achieved without any oscillation for both Wood and Berry model and VA model. Figures 1 and 3 show the simulation studies for the obtained controller with good tracking of the servo response for WB model as well as VA model, the response shows that the controller provides proper tracking with the given setpoint.

5 Future Work

The above-presented algorithm is to design an PI controller based on the parameters of loop transfer function for a square MIMO system. This methodology can be extended to non-square MIMO systems, where the two tray temperatures and pressure is considered as the process variables.

Acknowledgements The authors would like to express their deepest appreciation to the Dept. of Electronics and Instrumentation Engineering, Sri Chandrasekharendra Saraswathi Viswa Maha Vidyalaya, Kanchipuram for providing workspace for doing the simulation studies and Dept. of Instrumentation and Control Engineering, Manipal Institute of Technology, Manipal, Karnataka, for providing the facility to perform the real-time experimental work.

References

- Chen, Q., Luan, X., Liu, F.: Analytical design of centralized PI controllers for high dimensional multivariable systems. In: Proceedings of the 10th International Federation of Automatic Control (IFAC) Symposium on Dynamic and Control of Process Systems. Mumbai, India, pp. 643–648 (2013)

2. Bhat, V.S., Thirunavukkarasu, I., Shanmuga Priya, S.: An experimental study on implementation of centralized PI control techniques on pilot plant binary distillation column. *Int. J. Chem. Tech. Res.* **9**(11), 244–251 (2016)
3. Bhat, V.S., Thirunavukkarasu, I., Janani, R.: Design and implementation of MSC based multi-loop PID controller for a pilot plant binary distillation column. In: IEEE International Conference on Circuits Power and Computing Technologies (2017)
4. Luan, X., Chen, Q., Liu, F.: Centralized PI control for high dimensional multivariable systems based on equivalent transfer function. *ISA Trans.* **53**(5), 1554–1561 (2014)
5. Maghade, D.K., Patre, B.M.: Decentralized PI/PID controllers based on gain and phase margin specifications for TITO processes. *ISA Trans.* 550–558 (2012)
6. Vu, T.N.L., Lee, M.: Independent design of multi-loop PI/PID controllers for multi-delay processes. *World Acad. Sci. Eng. Technol.* **60**, 703–708 (2009)

A Hypothesis on Ideal Artificial Intelligence and Associated Wrong Implications



Kunal Shah, Pradeep Laxkar and Prasun Chakrabarti

Abstract The paper points out certain novel axioms viz., Selflessness, Humility with Ambiguity and Reinforced learning. The axioms have been correlated with the fact towards realising ideal Artificial Intelligence and the associated wrong implications.

Keywords Selflessness · Humility · Ambiguity · Reinforced learning · Artificial intelligence

1 Non-conventional View on Artificial Intelligence

At the forefront of all the problems faced by the development of AI is the fact that the industry leaders today are either ignorant towards or uninformed of what AI is supposed to be. The vanguard of the industry has repetitively tried to mirror human beings in their programs and products, knowing fully well that what they're trying is little more than an expensive gimmick. It's widely acknowledged today that the full potential of AI is not tapped yet, primarily because finding funding for legitimate research in the science is hard to come by, whilst simultaneously precious resources and funding are made available for gimmicks like Hanson Robotics' *Sophia*. This machine is a glorified 'question answering' machine that uses the bleeding edge of technology with AI to fuel man's fantasy of creationism. Concurrently, IBM with 'Watson' and its determination to duplicate human nature was so blinded by their objective that it cost the company 5 years of declining revenue before the project finally stopped getting funding. The sheer monetary cost for research in the project standing at \$1.8 Billion for 3 years—the expectations from which were sky high—only for the final outcome to be a disappointment, a waste of precious brain power and more importantly money. IBM pushed Watson into all sections of commerce from recipes to health care, clearly unsure of what its product was actually capable of or what its purpose was supposed to be.

K. Shah (✉) · P. Laxkar · P. Chakrabarti

Department of Computer Science Engineering, ITM Universe, Vadodara 391510, Gujarat, India
e-mail: kunalshah10011998@gmail.com

This further reinforces the idea of misuse of capital by nurturing research in point-less products and avenues for a largely versatile piece of underlying technology. Humanity's primitive understanding of intelligence is evident in its attempt to replicate said intelligence synthetically in AI, IBM's attempt with *Watson*, and Hanson Robotics' *Sophia*, or more accessible attempts like *Siri*, or *Bixby*, which are popular digital assistants on smartphones, to name a few. These systems are fundamentally flawed as they seem far too artificial and one cannot integrate them in their daily lives smoothly. Another big problem with these devices is that they cannot be scaled and their functionality is also very limited. The programs and systems stated above serve as a good example to highlight the bigger problem with AI today, these systems try to eliminate humans from the equation or try to control very superficial aspects of life; there is no deep impact. Essentially these systems lack intuition, something that is implied for a system that is labelled 'Artificially intelligent'.

In relatively older inventions, it's of note that single cognitive features were dialled to the maximum, like the calculator is vastly better at arithmetic and logical reasoning problems than humans. A GPS has superior spatial cognition than a human, and search engines have superior long-term memory when compared to humans, and all these devices are fundamentally in the right step, trying to augment man's cognition wherever it lags behind technology. Embedding minds with this technology has helped man make countless more discoveries and paved the way for the digital age.

2 Characteristics to Better Understand AI

There are three characteristics that hold key importance for a better understanding of AI and reducing all the negative connotations laymen and scholars alike link to it due to popular culture and ignorance. This improved understanding will lead them to embrace the idea of superintelligence, which in turn will create a domino effect in the field, leading towards the development of a safe and human-compatible AI.

2.1 Nascent Notion of Intelligence

Human kind's primitive understanding of what intelligence is can be encapsulated by what *René Descartes* very famously said, '*Cogito ergo sum*', which translates to 'I think therefore I am'. This phrase gave way to the idea of having a conscience, which paved the way for computational expert systems. This quote is from the eighteenth century, and since then, there has been no real progress in answering the question of, why humans have a conscience, or other pertinent questions like, do other species have a conscience? What constitutes for a conscience? Even still man has been on a crusade to synthetically replicate something he cannot define. This is not the only aspect of intelligence that humans got wrong, for example, people perceive intelligence as a one-dimensional thing, like IQ points. One completely discounts the role

of other cognitive features such as self-awareness, learning, emotional knowledge, deductive reasoning, planning, creativity, spatial awareness and many more. Think about intelligence as an orchestra, where each of these cognitive features is playing a different note, one louder than the other and the intricate symphony that these notes create is what one must refer to as intelligence.

The ideal AI's fore ordinance must be to augment human cognition. The wheel per se was invented so that it could augment the transportation capacity of humans, when used right, the wheel not only served its purpose, but laid foundation to age of curiosity, where one after another, man made discoveries and invented countless tools to pave the way for the world mankind lives in today. A more age-appropriate example could be the fact that the commodore is the most efficient animal on the planet, but given a bicycle, man becomes the most efficient animal in the world, now imagine a human mind having the most advanced AI in history at its command, now imagine every mind everywhere having access to it, the human race would collectively witness the biggest quantum leap in terms of democratised knowledge and information ever in recorded history.

Hence, AI's inherent fore ordinance must be to augment cognizance, any other use of that technology is short-sighted and enormously pointless in comparison to problems that scientists and thinkers could only define over the years but not answer. What happens beyond the event horizon of the black hole? What is dark energy? Are humans alone in the universe? These questions could finally have answers and for the first time in recorded history, man could catch up with his curiosity itself.

2.2 Comprehension of the True Potential of A.I

The potential implications of the invention of an ideal AI are in simple words, world changing. The human race could witness a Second Industrial Revolution and this time on a scale that will touch humans far more instantly than the last one. The First Industrial Revolution was as impactful as it was because man's endeavour or creating 'Artificial energy' was successful. Engineers could suddenly create adequate power from fossil fuels and this discovery was the effective juggernaut that propelled all industries towards mechanising. The modern agricultural revolution was also a by-product of this capacity to suddenly cultivate bigger lands more efficiently, innovators built roads, skyscrapers, cars, all these inventions and developments across all sectors were attributed man's capacity to generate artificial energy. In the same vein, AI will be at the centre of development in all sectors of society and industry alike. If one thinks about it, this shift in the paradigm is already here, human live in smart homes which in turn are in smart cities, most use smartphones, drive smart cars, and people are connected to the digital bubble now more so than ever before. The astounding thing is that developers barely have applied AI to its maximum potential in any of these scenarios, but these 'Smart' (i.e. devices that integrate AI in any way, shape or form in addition to but not limited by devices that augment human cognition in any way) devices have already altered the modern way of life significantly.

Scientists will soon be working with these systems of AI, which tackle problems that the human race did not face before. A case in point being climate change. A man could barely keep up with weather prediction models even with state-of-the-art computers and yet, the sheer uncertainty of nature made it a near impossible task for meteorologists to make the right prediction every time. With augmented cognizance, humans could certainly make more of an impact on the problem they face with climate change. These systems are essentially like ‘Exo-suits’ for the mind; they augment the mind’s capacity to think in ways it could biologically not, helping it explore solutions to problems which could not have arrived to as normal humans.

Adding intuition to all these systems will push technology over the inflection point in the graph of development and sophistication and usher in an age of Augmented Cognizance that quite frankly humanity cannot describe because in this age the very act being human will vastly differ from what its connotation is today.

3 Disproving Widely Held Misconceptions About AI in the Scientific Community

The question, ‘Is A.I. something to fear?’ has piqued the interest of enthusiasts and critics of AI alike ever since its inception. Every innovation attracts upon itself some trepidations. History is witness to this, upon the invention of the television set, the scholars petitioned that the layman would procrastinate and become bovine in nature. When the idea of e-mail was conceived, the social order lamented at the lack of formality this invention brought, and the intimate and personal touch of the letter could not be replaced. As the internet became an omnipresent fixture of human lives, it was rational to suggest that man could over time lose the aptitude to memorize things. It is fair to suggest that there’s verity in all these assertions, but the fact of the matter is that these innovations delineate mankind’s way of contemporary life, where the idea of information exchange has been taken for granted irrespective of the medium of propagation, this enriched the human experience in considerable ways. ‘2001: A Space Odyssey’ a film by director Stanley Kubrick from 1967 anthropomorphizes all stereotypes that mankind has come to associate with AI, in the film HAL 9000 is a sentient and conscious computer programmed to support the discovery space crew on their journey. Hal on the journey chose to value the mission goals over the human lives of the space crew. Whilst the character of HAL is fictional, it echoes humankind’s fear of being enslaved by a being of superintelligence who’s is indifferent to humanity.

The progress in the field today made by researchers and scientists is very much along the lines of HAL, it’s here where with all due respect everyone in the technology business collectively forgot what the fore ordinance of this technology was, to improve human capacity. Instead, the industry headed in the direction of monetising a technology that it clearly did not understand fully. Instead, the collective effort needs to be focussed around trying to create the ‘Ideal A.I.’ where the system pursues its

goals with nuance and not with a single-minded approach. If and when this objective is achieved, man would be a step closer to a human-compatible AI.

As per [1] the book ‘Superintelligence’ by renowned academic Nick Bostrom, in which he postulates that AI’s very existence induces an existential crisis for humanity, whereby its very existence is foreshadowing the obsolescence of humanity itself. Dr. Bostrom draws this conclusion after taking into consideration that cognitive systems absorb data quantum timescales, meaning an AI will have inferred and interpreted all forms of information produced by man, books, texts, papers, videography, audio files, etc., in its quest to expand its knowledge base, this scenario makes an allusion to the AI’s gluttonous hunger for data and eventually that day might come where the AI establishes the goals set for it by its human creators no longer line up with the objectives the AI sets for itself.

Whilst Dr. Bostrom is held in high regard by his peers in the scientific community and with all due respect to him and his peers, this paper endeavours to disqualify the fears and objections held by Dr. Bostrom to an extent. A reference to Hal can be drawn over here again, Hal became a problem for the protagonists in the film only because it was in charge of all of the spacecraft’s mission-critical systems, the point over here being in the real world no system ever has absolute autonomy over anything.

Engineers and safety experts always build in fail-safes for all their creations, there exists a whole sub science dedicated to this exercise called safety engineering, because when the technology fails, human intervention can still solve a problem and prevent a calamity from happening. Dr. Bostrom failed to acknowledge this obvious real-world adaptation of the AI technology, to further elaborate on this point in the real world assuming that the AI does become omnipotent and decides to face off with humans, it would still have to compete with us for resources. AI at that point would directly compete with all of the world’s biggest economies. This is where mankind has the upper hand, the worst-case scenario for us with AI has a simple solution, all man needs to do is starve it of the resources that are critical to its functionality.

Additionally if the three maxims stated above are used as the building blocks in the design of this AI, then the stigmas and misconceptions about AI will be nothing more than mere misunderstandings, instead mankind might witness the opposite where humanity learns to cherish and use AI to its maximum potential, and it could champion all endeavours of progress across all fields of science and commerce alike. An AI built with human morals and values will most certainly give mankind the best shot at a human-compatible AI, whilst this is a fair deduction of the scenario in the real world one cannot with certainty shield the human race and the planet from all arbitrary acts of violence and calamity; humans are flawed and fickle, and the outliers amongst humans are extremists, and AI in the wrong hands could be used as a weapon of mass destruction, but again this is not possible, simply because the maxims are designed to value human life, all human life.

Furthermore to stave off any such developments in the field Max Tegmark the author of ‘Future of life foundation’, has designed along with other leaders in the field the ‘Asilomar Principles’ to stigmatize the research in AI’s potential as a weapon. These principles were developed by the van guard of AI research community, this

makes the underlying technology extremely difficult to be utilised as a weapon, since the people who designed it, will not give it that capacity or free reign to develop required skill set.

4 Proposed Novel Axioms to Design the Ideal Artificial Intelligence

To program and design a human-compatible AI system, researchers need to mimic in these synthetic systems, a sense of just how whimsically humans think. Whilst this seems to be an engineering problem that's far too difficult, its solution can be broken down to three simple axioms:

- The first Axiom: Selflessness;
- The second axiom: Humility with Ambiguity;
- The third axiom: Reinforced learning.

4.1 *Selflessness*

In an ideal AI, the creator's primary objective is to ensure that the AI comprehends the objectives set to it by us humans, and to do this, the AI will need to learn and adopt all human values. In this instance, the word 'value' entails all sets of specific values, moral and emotional, extrinsic and intrinsic and man would have to factor all other kinds of intelligence too as stated earlier in the text. This is necessary as human nature is represented in the light of fuzzy logic, and that reflects in human ethics and behaviour. As per [2], a sense of altruism essentially as a core principal in AI partially infringes upon Asimov's law that conditions 'robots have a sense of self-preservation' but realigning the AI's core value set prioritises its maker's existence of over its own, irrespective of the circumstances, and this realignment, in turn, reinforces the third law determined by Isaac Asimov. This makes the fear of 'Intelligent machines taking over the world null and void'.

4.2 *Humility with Ambiguity*

This maxim is important because it gives fundamental context to the AI in terms of the values developers want it to adopt. Since AI is synthetic in nature and it cannot understand what human values are, it would maximises them, all whilst not knowing what it is that these values are for. The inherent ambiguity in human values is of the creator's gain over here, the problem of steadfastness in an AI when it is pursuant of its objective can be avoided, like HAL. In order for AI to be relevant, it has to have a

coarse idea of what the user wants it to do. It acquires this data largely by observing humans, and by extension human choices. The question one might pose here is ‘What happens if the AI is ambiguous about its aim?’ The answer is rather simple, the AI would fundamentally reason differently when compared to its predecessors. This iteration of AI factors in the situation where a man would choose to ‘flip the switch’ and turn it off, but only if the AI is inherently doing something wrong. Since AI has no concept of what wrong is, but it would choose to not do it. On closer observation, it can be discerned that the first two maxims are in play over here. Which means the AI would let the humans turn it off. Simple math statistically indicates the approximate motivation an AI has to licence the user to turn it off, and this tendency of the AI is proportional to the degree of vagueness of the objective set for the AI. Below stated is an equation that embodies the mathematical equivalence of axioms 1 and 2.

$$P_i + (-P_i) = 1 \quad (1)$$

where $i = 1$ to n , where n denotes all the values morals and associated nuances.

P_i Denotes the probability of these values, morals and associated nuances being used in the AI’s decision theory.

$-P_i$ Denotes the probability of all the misconceptions, incomplete cognitive features and the primitive design structure used to make AI.

4.3 Reinforced Learning

This principal is important as it comes into effect after the system is turned off. The AI infers that there is something wrong in its approach to achieve the objectives it should be pursuing, because the action of it being terminated signals that the actions taken by the AI were wrong. An AI designed to learn from its errors is by far better than an AI that was created in any alternative way. The example stated above is a prime example to showcase what man ventures to undertake with human-compatible AI. This principal could potentially draw upon itself the ire of the scientific community, because man often behaves badly or in extreme cases, deplorably. A lot of this unsavoury behaviour it could be argued would be adopted by an AI that is built in the image of man using human values. Which means that inevitably the AI will become immoral, and corrupt itself just as man has himself. Except this is not the classical case of ‘The fruit of the poisonous tree’, the entire purpose of these dictums was to provide nuance for why mankind does what it does, and help it understand why it is that mankind does it. The ideal AI’s goal is not to overpower man, but empower him. To allow man the privilege to solve problems, he yet does not know exist. The figure given below embodies the novel idea to showcase reinforced learning using a Markov model (Fig. 1).

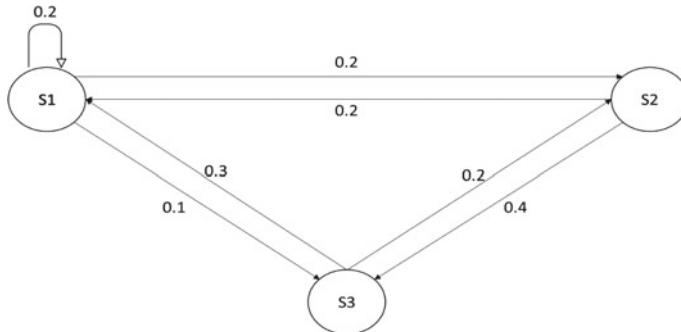


Fig. 1 Markov model to delineate a novel approach to showcase reinforced learning where S1, S2, S3 are all representatives of different conclusions that can be reached to for a particular problem by accounting for all the weights ‘P’ that probability of all the values, morals and associated nuances as programmed by developers

4.3.1 Markov Model’s Role in Reinforcement Learning

According to [3], the Markov model depicted is a Markov Decision Processes are intended to be a direct outline of the problem of learning for any agent from interaction to achieve a predefined objective. The agent and its environment interface constantly, where the role of the agent is to select actions, and the effects of said actions manipulate the environment, hence making it a dynamic process whereby the environment’s new state is an outcome of the action of the agent and the environment’s ever changing states result in subsequent actions that are taken by the agent to alter the environment to desired state. Strictly, MDPs are utilised to define problem spaces (i.e. environment) for reinforcement learning, wherein the problem space is totally observable.

To better comprehend what an MDP is, it’s of paramount importance to understand what the Markov principle fundamentally states, ‘*The future is independent of the past given the present*’. Which means to predict the weight of the action ‘P’ that is to be taken in case the environment has a state S_n is completely independent of what the environment was on S_{n-1} . This means that the present state S_n is an adequate figure which characterises the future states S_{n+1} in the given model just as accurately as a model that includes the history of all the states prior to S_n . In mathematical terms, a state S_n has the Markov property, if and only if;

$$P[S_{n+1}|S_n] = P[S_{n+1}|S_1, \dots, S_n] \quad (2)$$

The state transition probability function for interstate transition in the Markov model is

$$P_{SS'} = P[S_{n+1} = S'|S_n = S] \quad (3)$$

The equation stated above signifies a probability distribution as a function of ‘P’ over next possible successor states ‘S’, given that the present state is S_n . (i.e. the agent is in some state, there is a probability P_1 to go to state S_1 , and another probability P_2 to go to state S_2 and so on.).

This transition function can be represented in a matrix, where all transactions to and from one state must = 1.

$$\begin{matrix} P_1 & \dots & P_{1n} \\ \vdots & & \vdots \\ P_{n1} & \dots & P_{nn} \end{matrix} \quad (4)$$

4.3.2 Markov Reward Process’s Role in Value Loading

Assuming that scientists and programmers alike have consensus to fashion an AI on the principles stated above, another major obstacle remains. How is it that they load these values, morals, and the nuances that go along with them? The answer to this question is of critical importance to the development of AI itself. A probable solution to this problem could lie in the *Markov Reward Process* (MRP).

According to [4], MRP is a method that utilises value decision, determining the amount of reward accrued via a specific sequence that is sampled. An MRP is represented by the following notations of (S, P, R, γ) where P is a state transition probability function, S is the state space and R denotes the reward function, which signifies the instant reward the agent must anticipate from state S at the given particular moment. γ signifies the discount factor, which lies between $[0, 1]$. Where if the value of γ closer to 0, it means that the agent is short-sighted in terms of rewards meaning only immediate rewards interest the algorithm, but if γ is closer to 1, it means that the algorithm is long-sighted and it is pursuing a goal that gives rewards in the long term. Here, we establish that γ as a virtual horizon, which denotes with its value if a particular objective is worth pursuing by comparing its reward value to its alternatives.

$$R_s = [R_{n+1}|S_n = S] \quad (5)$$

$$G_n = R_{n+1} + \gamma R_{n+2} + \gamma^2 R_{n+3} + \dots = \sum_{k=0}^{\infty} (\gamma^k R_{n+k+1}) \quad (6)$$

Here, G_n signifies the summation of all the rewards that are accrued from time t .

From Eq. 6, we derive the state-value function of the MRP starting from S . This equation is the Bellman Equation. The Bellman equation is used to find the optimal value function

$$v(s) = [G_n|S_n = S] \quad (7)$$

According to [5], the value of S signifies the reward the agent accrues when exiting the state, plus the average discounted reward as a function of γ over possible subsequent states, where the probability of landing on a possible subsequent state is multiplied with the value of each possible successor state in Markov chain.

$$v(s) = R_s + \gamma \sum_{s' \in S} (P_{ss'} V(s')) \quad (8)$$

Since Eq. 8. is linear in nature, we simplify it to

$$V = (I - \gamma P)^{-1} R \quad (9)$$

The computational complexity of this equation is $O(n^3)$, for n states.

The beauty of this method is we could help the ‘Ideal A.I.’ absorb all the values, morals and ideals by programming the internal incentive system of the AI by using the Markov Reward Process such that it chooses to follow ethics and behavioural patterns by identifying them using the three truisms stated above. In this way, the complex and tedious problem of Value Loading in AI can be tackled. Building AI this way is practically possible, deep learning science has progressed vastly in the last decade, AlphaGo from Google’s deep mind project being a prime example. Construction of a cognitive AI system is profoundly different than programming a software-exhaustive system. Cognitive systems do not need human programming to improve their functionality, they recursively learn from their mistakes and alter their knowledge base and correct their value set to improve their efficiency in pursuing the goals set to it. Consider the game playing AI for chess, human programmers first handcraft the rules of the game for the system, then simulate thousands upon thousands games, where the AI learns to distinguish a good game and a bad game. The same logic applies for an AI medical assistant where experts craft into the knowledge base all the information available on cardiology, symptoms, conditions, disorders, but with that, experts teach the AI all the complications and false positives of the underlying symptoms. Scientifically, this is called the **Ground Truth**. This means teaching AI a semblance of what being human is and how morals and values influence it as a being. Thus giving it much required context for man’s complex and non-uniform behaviour and distinctly non-binary values.

5 An Ideal AI’s Application as a Generative Design Tool

Every tool that man has ever fabricated, cultivated or programmed has one thing in common, they are all passive tools. These tools can only do what the user does with them and nothing more. For example, knives only cut where they are struck, the pen only writes what the hand commands it to, and even the most advanced computers have to be programmed manually, and every end user on those computers can write

programs that the compiler in the computer can understand. This has been the case since the dawn of man's progress, where humankind manually compelled its tools to bow to its will, to most this seems normal, (as per Kazi et al. [6]) but innovator Maurice Conti and his generative design tool 'Dreamcatcher' challenge these norms. Generative design tools use algorithms to fuse physical geometry with theoretical efficiency to birth new designs autonomously.

What Conti and his team of researchers have accomplished with 'Dreamcatcher', is quite remarkable, where the users only give the system a very generic and plain goal and configure some constraints, what happens there is rather game-changing in nature, where AI instinctively starts to find a viable resolution in the solution space. These algorithms work very much like evolution itself, bettering themselves with every iteration and refining the solution space with better and more decisive results. For example, take the assumption that a perfect car chassis is to be designed. Dreamcatcher in this scenario would require raw perceptual data points from an existing car chassis, all the forces and conditions that it is subject to. Once data scientists have accumulated a data set that is representative of the best, average and worst case scenarios faced by the chassis. Dreamcatcher processes this data to create a chassis that meets all the preset criterions, and devises a design that could not have been imagined by a human engineer, and it does this all without any human design input, there were never any drawings, and the system had to start from scratch. The result is a chassis that is vastly superior to any fabricated in all of motoring history. Airbus, the airline manufacturers have a project for an airliner that is in development which will completely be developed by AI, the fact is that this shift towards generative tools is already underway, and what makes this change remarkable is the fact that these are first-generation generative tools, they have glaring flaws and problems, yet the results they yield are by far better than any produced by the most skilled humans, in most niche fields this gives innovators and inventors a glimpse of the future where these systems will integrate with humans to create more efficient, more environmentally friendly, and objectively better designs than any human designers.

6 Conclusion

The paper entails a novel proposal of three axioms that definitely make a breakthrough in modern Artificial Intelligence research. The hypotheses have been supported by proper justifications.

References

1. Bostrom, N.: Superintelligence: Paths, Dangers, Strategies. Oxford University Press, Sept 3rd 2014. Accessed 9 Dec 2017
2. Anderson, S.L.: Asimov's "three laws of robotics" and machine metaethics. *AI & Soc.* **22**(4), 477–493 (2008)

3. Gouberman A., Siegle M.: Markov reward models and markov decision processes in discrete and continuous time: performance evaluation and optimization. In: Remke, A., Stoelinga, M. (eds.) Stochastic Model Checking. Rigorous Dependability Analysis Using Model Checking Techniques for Stochastic Systems. ROCKS 2012. Lecture Notes in Computer Science, vol. 8453. Springer, Berlin, Heidelberg (2014)
4. Bradtke, S.J., Duff, M.O.: Reinforcement learning methods for continuous-time Markov decision problems. In Advances in Neural Information Processing Systems, pp. 393–400 (1995)
5. Baird, L.: Residual algorithms: Reinforcement learning with function approximation. In: Machine Learning Proceedings, pp. 30–37 (1995)
6. Kazi, R.H., Grossman, T., Cheong, H., Hashemi, A., Fitzmaurice, G.W.: DreamSketch: early stage 3D design explorations with sketching and generative design. In: UIST 2017 Conference proceedings: ACM Symposium on User Interface Software & Technology (2017)

Quantum Communication Based on the Photon Area Function as Basis States for Photon Qubit



Vineet Kumar

Abstract In the field of communication either by classically or by quantumly light is the natural carrier of information through both material space and vacuum therefore it is essential to look upon the different aspects, here, the mensuration and the geometry chapters of photon were considered as this may aid in understanding of different features and applications in enhanced way. The proposed mensuration and geometry chapters of photon deal with photon area function to serve as basis states for photon qubit under peculiar conditions. Based on the photon area function for geometry of circular type in perpendicular to the line of travel as basis states for photon-based qubit to serve the purpose of secure quantum communication, the complex coefficient so establish in this paper are like any ordinary function to which the calculus approach of different, for instance, differentiation, integration, etc., can apply.

Keywords Photon qubit · Photon area function · Quantum hole · Quantum communication · Travel line

1 Introduction

Photon (named by G. N. Lewis in 1926 [1] for the quanta of light a concept introduced by M. Planck in 1900 [2, 3] involves integral replacement $\int dv$ by discrete sum $\sum_i v_i$ for different frequencies contained in the black body) localize both in space and time reported to different features and applications like quantum no-cloning [4], quantum no-deletion [5], quantum cryptography [6, 7], etc. Photon

- with associated spatial length λ of order 10^{-9} m to it; and
- with travelling speed c of order 10^8 m/s to it,

on a time period T of order 10^{-17} s in satisfaction with relation $T = v^{-1} = \lambda/c$ interact weakly with their surroundings even over a long distance, due to for each

V. Kumar (✉)

Department of Electrical Engineering, Kurukshetra University, Kurukshetra 136119, India
e-mail: vineet05k@gmail.com

1 m there is a holding of order 10^{-8} s, as a consequence this results lower noise and loss. ν is the frequency. Out of the different mathematical chapters in concern of photon here from different aspects of a photon book, for instance:

- i. *algebraic chapter*—based on energy $E = h\nu$, momentum $p = h/\lambda$, etc., description of photon along with their conservation principles and other related equations, where h is Planck constant having value $6.6260755 \times 10^{-34}$ J s;
- ii. *probabilistic chapter*—relates with exhaustive, mutually exclusive and equally likely description of photon based on the set of measurements as events for related photon variables;
- iii. *geometric chapter*—based on the geometry description in relation to photon perpendicular to the line of travel such that it does not limit to a point (or 0-dimensional object) as ‘the point is that which has no part [8]’ according to Euclid mention in his book element part 1 and therefore no action consequently. In case of surface and volume, the geometrical object, by the case of collection of points, has the least two and three-dimensional space requirement for it to determine, which in correspondent by other way indicates the consideration of least in terms of optionality over space to exist. The background potential in case of two-dimensional geometrical object to bagged up it is relatively more than that of three-dimensional geometrical object, but less than that of one-dimensional object. Next;
- iv. *mensuration chapter*—based on the length, area, etc., description in relation to a photon due to the geometry of peculiar around the line of travel;

the mensuration and geometry chapters of photon needed to give more stress on and is the main objective of this paper, as in the continuum state of travel, these two chapters not only helps in understanding the manner of dealing of photon with elementary particles but the application of different as well can also be realized in a better way. In this paper, for secure quantum communication, the photon area function is suggested to serve as basis states for photon qubit under peculiar conditions. The mensuration and geometry-based concept for photon introduce first in quantum hole [9] paper as photon area function where it conclude that the optical hole at microscopic level is not a void place but a quantum object.

2 Probabilistic Chapter of Photon for ‘Particle Like’ and ‘Wave Like’ Nature

Based on the set of variables measurement; for instance, position, momentum, energy, etc.; as events for photon, there are two exhaustive cases for it, where one case corresponds the wave nature while other is the particle nature. In other words, if one discusses about the nature of outcomes based on the measurement of photon variables, then there are two exhaustive cases only either as of wave or as of particle, there is no third case like the no third case other than head and tail outcomes in coin tossing experiments. The measurement of a photon variable, at a time, indicate

it either as of wave or as of particle means the occurrence of one of the nature precludes the occurrence of all other, therefore such basis of outcomes for photon variables measurement are regard as of mutually exclusive type. In other words, if one when measuring the photon variable, then the behaviour of only one from such as of wave or as of particle conclude at a time only. For instance, measuring the position indicates ‘particle like’ nature while measuring the momentum indicates ‘wave like’ nature [10]. In both cases of measuring the position and momentum of photon, at a time, indicates only the occurrence of ‘particle like’ and ‘wave like’ nature at a time and therefore precludes the occurrence of ‘wave like’ and ‘particle like’ nature, respectively. This particle-like position and wave-like momentum of photon related by h as $\Delta x \cdot \Delta p_x \approx h$ [11], which with $\Delta x = c\Delta t$ condition gives $\Delta t \cdot \Delta(p_x) \approx h$. As the unit of cp_x is energy, let $cp_x = E$, then position–momentum relation modify to become as time–energy relation given by $\Delta t \cdot \Delta E \approx h$. On equating the constant h , that marks the limit, for both position–momentum relation and time–energy relation, it gives $\Delta x \cdot \Delta p_x \approx \Delta t \cdot \Delta E$, which on rearranging gives Eq. (1) having a unit of force.

$$\frac{\Delta p_x}{\Delta t} \approx \frac{\Delta E}{\Delta x} \quad (1)$$

The position–momentum relation and time–energy relation for photon over constants h/c and hc with $c = 2.99792458 \times 10^8$ m/s also become as Eq. (2) and Eq. (3), respectively. Equation (1) also gives the relation of $\Delta E \approx c \cdot \Delta p_x$.

$$\Delta t \cdot \Delta p_x \approx \frac{h}{c} \quad (2)$$

$$\text{and, } \Delta x \cdot \Delta E \approx hc \quad (3)$$

Other than position–momentum relation and time–energy relation, all of equations from (1) to (3) so obtained here determines that the particle-like and the wave-like nature of photon are equally probable, therefore such basis of outcomes for photon variable measurement are regarded as of equally likely type. In collective manner, the wave and particle duality of photon determines purely exhaustive, mutually exclusive and equally likely type.

3 Mensuration and Geometry Chapters of Photon for Photon Area Function

With the time complex exponential term of electric vector E [12] as $L(\lambda, \varphi, t) = e^{i2\pi ct/\lambda} \cdot e^{i\varphi}$ such that $L(\lambda, \varphi, t) = L(\lambda, \varphi, t + r\lambda/2c)$ where r is integer and φ as the phase constant, if the spatial complex exponential term, suppose along z -axis as $e^{i2\pi z/\lambda}$, introduce, then it gives the decomposed electric field of light as $|E| \propto$

$X \cos \omega t + Y \sin \omega t$, where $\omega = 2\pi c/\lambda$, from $M(\lambda, \varphi, z, t) = L(\lambda, \varphi, t)e^{i2\pi z/\lambda}$. This $M(\lambda, \varphi, z, t)$ is the space-time complex exponential term of electric vector \mathbf{E} . The decompose electric field of light over classical quadratures X and Y as function of k and φ with $k = 2\pi/\lambda$ play the role in quantum mechanics as suggested by B. Julsgaard et al. [13], which satisfy the canonical commutation relation $[\hat{X}_L, \hat{P}_L] = i$ with \hat{X}_L and \hat{P}_L as two quadrature phase operators. i is the imaginary symbol. Other than $M(\lambda, \varphi, z, t)$ for \mathbf{E} , the space-time complex exponential term for magnetic vector \mathbf{H} let as $N(\lambda, \varphi, z, t) = e^{\pm i(\omega t - kz)}e^{i\varphi}$ also need to be consider on as complementary which on collective after the solution from $\nabla^2(\mathbf{E}_s/\mathbf{H}_s) - \gamma^2(\mathbf{E}_s/\mathbf{H}_s) = 0$ determines that the light is the manisfestation of electromagnetic wave [14, 15] as the speed so evolve is equal to the speed of light and also due to property satisfaction. γ^2 is a complex number on individual basis for each \mathbf{E} and \mathbf{H} equations.

Light characteristically as ‘electromagnetic’ [16] also allowed to be suppose for the single photon as well on the ground of involve mensuration and geometry with it, where the area function being in perpendicular to the line of travel in continuum is $a(\lambda, d, x, y, z, t)$ (or simply $a(\lambda, d, z, t) = a(\lambda, z, t)$ along z -axis). The variable $d = d(\lambda, z, t)$ of the photon area function $a(\lambda, d, z, t)$ for circular geometry is the diameter of it with satisfaction of partial differential equation of $(\partial^2/\partial z^2 - \rho^2 \partial^2/\partial t^2)d(\lambda, z, t) = 0$ that leads to ordinary differential equations of $(\partial^2/\partial z^2 - \rho_1^2)d_1(\lambda, z) = 0$ & $(\partial^2/\partial t^2 - \rho_2^2)d_2(\lambda, t) = 0$, and therefore determines the same space-time complex exponential term as that for \mathbf{E} and as that for \mathbf{H} as $e^{\pm i(\omega t - kz)}e^{i\varphi}$. ρ_1^2 and ρ_2^2 are constants. This diameter function of the photon $d(\lambda, z, t) = d_1(\lambda, z).d_2(\lambda, t)$ indicates the transverse one-dimensional wave nature of it. So like the diameter function $d(\lambda, z, t) = d(\omega, k, z, t) \equiv d_o e^{\pm i(\omega t - kz)}$, the area function $a(\lambda, d, t)$ of photon can also be given as $a(\lambda, d, z, t) = a(\lambda, z, t) = a(\omega, k, z, t) \equiv a_o e^{\pm i(\omega t - kz)}$ where a_o is the area amplitude in correspondent to that of diameter amplitude d_o . The photon with a case of action around the travel line is possible only if the photon does not involve a point condition as the point has no part and therefore no action. The photon area function $a(\omega, k, z, t)$ in protrudation over differentials $\partial/\partial t$ and $\partial/\partial z$ for moduli ratio is given as

$$\left| \frac{\partial/\partial t a(\omega, k, z, t)}{\partial/\partial z a(\omega, k, z, t)} \right| = c \quad (4)$$

The constant c of Eq. (4) also holds for $\left| \frac{\partial/\partial t d(\omega, k, z, t)}{\partial/\partial z d(\omega, k, z, t)} \right|$ as well.

4 Photon-Based Qubit Representation over Photon Area Function as Basis States

The art of (*en* – & *de*–)coding information is the cryptography, which have been existed for a long time from the past for secure communication. The quantum

communication (or q -communication) is now being the centre of research area in the field of information technology, today. Most of the countries worldwide taken q -communication system into serious and consequently calls for different scheme/proposal as well. According to Jane Qiu [17], China starts installing the world's longest q -communication network, which includes a 2,000 km link between Beijing and Shanghai. In addition, a team led by Hayford together with ID Quantique has started installing a 650 km link between Battelles headquarters and its offices in Washington DC.

So, what is the difference between classical communication (or c -communication) and q -communication? and how q -communication is better than c -communication? In c -communication, information is encoded in amplitudes of light pulse, which are then detected, converted into electric signals and stored as electric charges or magnetization of memory cells while in q -communication, information is encoded in quantum states but recorded by the same classical means [13]. q -communication plays a vital role in military use, national security and in everyday life due to a novel way of communication with unconditional security [18].

In consideration with mutually exclusive events for a transistor [19] 'on' and 'off' denoted by $|O\rangle$ and $|F\rangle$, respectively, with parameters $\alpha \equiv \alpha(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)$ and $\beta \equiv \beta(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)$ over set of variables $\bigvee_{l=0}^n x_l = (x_0, \dots, x_n) = (t - x, y, z, \dots)$ and $\bigvee_{l=0}^m y_l = (y_0, \dots, y_m) = (\omega, k, \varphi, \dots)$ as complex-valued \mathbb{C} coefficients for it, then, when employed for information encoding:

- i. On the ground of classical set (or crisp element) theory in correspondent, $\alpha(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)$ holds 1 or 0 while $\beta(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)$ holds 0 or 1 complementarily. This type of approach came under conventional communication system and therefore term as c -communication system where 0 and 1 are regarded as bit and is the basic unit of information.
- ii. On the ground of fuzzy set (or membership function) theory in correspondent, $\alpha(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)$ and $\beta(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)$ like that for c -communication does not involve solely in q -communication but is collectively defined by $\alpha(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)|O\rangle + \beta(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)|F\rangle$ let denote by $|\phi\rangle$ such that $|\alpha(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)|^2 + |\beta(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)|^2 = 1$. $|\alpha(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)|^2$ is the measuring of the state $|O\rangle$ and $|\beta(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)|^2$ is the measuring of the state $|F\rangle$. The qubit is the elementary of quantum system.

In general, the state $|\phi\rangle = |\phi(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)\rangle$ of a system used for information (*en* – & *de*–) coding based on superposition principle is described by Eq. (5) as

$$\left| \phi\left(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l \right) \right\rangle = \alpha \left(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l \right) |O\rangle + \beta \left(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l \right) |F\rangle \quad (5)$$

$\alpha(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l) \in \mathbb{C}$ and $\beta(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l) \in \mathbb{C}$. Basis states $|O\rangle$ and $|F\rangle$ of Eq. (5) here relate with photon area function carrying in $x - y$ plane and travel line along z -axis, such that it corresponds to the two straight line equations on variable z and t as $|O\rangle \rightarrow f(zt)$ and $|F\rangle \rightarrow g(zt)$. Basis states $f(zt)$ and $g(zt)$ over

complex exponential term $e^{\pm i(\omega t - kz)}$ determines $f(zt) \equiv \omega t - kz = p\pi$ and $g(zt) \equiv \omega t - kz = (p + \frac{1}{2})\pi$. p is integer. Equation (5) in terms of $f(zt)$ and $g(zt)$ become as

$$\begin{aligned} \left| \phi \left(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l \right) \right\rangle &= \alpha \left(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l \right) |f(zt) \equiv \omega t - kz = p\pi\rangle \\ &\quad + \beta \left(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l \right) |g(zt) \equiv \omega t - kz = \left(p + \frac{1}{2} \right)\pi\rangle \quad (6) \\ \text{with } 0 \leq \left| \alpha \left(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l \right) \right|^2 &\leq 1 \quad \text{and} \quad 0 \leq \left| \beta \left(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l \right) \right|^2 \leq 1 \end{aligned}$$

Over length λ , like same vibration condition in string wave, the coefficients of Eq. (6) also satisfy the condition given by Eq. (7), which yield Eq. (8) as

$$\begin{aligned} \left| \alpha \left(\bigvee_{l=0}^{n-1} x_l, z, \bigvee_{l=0}^m y_l \right) \right|^2 + \left| \beta \left(\bigvee_{l=0}^{n-1} x_l, z, \bigvee_{l=0}^m y_l \right) \right|^2 &= \\ \left| \alpha \left(\bigvee_{l=0}^{n-1} x_l, z + \lambda, \bigvee_{l=0}^m y_l \right) \right|^2 + \left| \beta \left(\bigvee_{l=0}^{n-1} x_l, z + \lambda, \bigvee_{l=0}^m y_l \right) \right|^2 &= 1 \quad (7) \end{aligned}$$

$$\text{so that } \left| \phi \left(\bigvee_{l=0}^{n-1} x_l, z, \bigvee_{l=0}^m y_l \right) \right\rangle = \left| \phi \left(\bigvee_{l=0}^{n-1} x_l, z + \lambda, \bigvee_{l=0}^m y_l \right) \right\rangle \quad (8)$$

As α and β are complex, then by consideration of $\alpha = |\alpha| e^{i\gamma(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)}$ and $\beta = |\beta| e^{i\vartheta(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)}$ the state $|\phi(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)\rangle$ modified to become in polar form as $e^{i\gamma(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)} (|\alpha| |f\rangle + |\beta| e^{i\vartheta - \gamma(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)} |g\rangle)$, which by supposition of $|\alpha| = \cos\theta(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)/2$ and $|\beta| = \sin\theta(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)/2$ for $0 \leq \theta(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l) \leq \pi$ in satisfaction with provided condition, Eq. (6) becomes as

$$\begin{aligned} &\left| \phi \left(\theta \left(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l \right), \varphi \left(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l \right) \right) \right\rangle \\ &= e^{i\gamma \left(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l \right)} \left[\cos\theta \left(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l \right)/2 |f(zt) \equiv \omega t - kz = p\pi\rangle \right. \\ &\quad \left. + e^{i\varphi \left(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l \right)} \sin\theta \left(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l \right)/2 |g(zt) \equiv \omega t - kz = \left(p + \frac{1}{2} \right)\pi\rangle \right] \quad (9) \end{aligned}$$

Equation (9) like Eq. (6) is the required qubit representation over photon area function in terms of polar form, where φ on $(\bigvee_{l=0}^n x_l, \bigvee_{l=0}^m y_l)$ is equal to $\vartheta - \gamma$.

5 Conclusion

The consideration of photon qubit on the ground of photon area function as basis states under peculiar conditions, i.e. attainment of $\omega t - kz$ of $e^{\pm i(\omega t - kz)}$ to $p\pi$ and $(p + 1/2)\pi$, for secure quantum communication is the objective of this paper and is given by Eq. (6) in connection with conditions mentioned in Eq. (7). Equation (9) is the polar representation of Eq. (6). Besides photon qubit on the ground of photon area function basis states, Eqs. (1)–(3) are also obtained here, where each of which indicates that the particle-like and the wave-like nature of photon are equally probable.

Acknowledgements The author would like to thank the referees for their direct and indirect support, which helped to improve the paper considerably.

References

- Lewis, G.N.: The conservation of photons. *Nature* **118**, 874–875 (2981), 18 December (1926). <https://doi.org/10.1038/118874a0>
- Planck, M.: On an Improvement of Wien's Equation for the Spectrum (1900). www.ffn.ub.es/luisnavarro/nuevo_maletin/Planck%20%281900%29,%20Improvement%20of%20Wien%27s.pdf
- Planck, M.: On the Theory of the Energy Distribution Law of the Normal Spectrum (1900). [http://hermes.ffn.ub.es/luisnavarro/nuevo_maletin/Planck%20\(1900\),%20Distribution%20Law.pdf](http://hermes.ffn.ub.es/luisnavarro/nuevo_maletin/Planck%20(1900),%20Distribution%20Law.pdf)
- Wootters, W.K., Zurek, W.H.: A single quantum cannot be cloned. *Nature* **299**, 802–803, October (1982). <https://www.nature.com/articles/299802a0>
- Pati, A.K., Braunstein, S.L.: Impossibility of deleting an unknown quantum state. *Nature* **404**, 164–165, 9 March (2000). <http://www.hri.res.in/~akpati/patinature.pdf>
- Wiesner, S.: Conjugate coding. *ACM Dig. Lib.* **15**(1), 78–88 (1983)
- Bennett, C.H., Brassard, G.: Quantum cryptography: public key distribution and coin tossing. In: Proceedings of the International Conference on Computers, Systems and Signal Processing, pp. 175–179, December (1984), republished 2014 in Theoretical Computer Science, Elsevier. <http://dx.doi.org/10.1016/j.tcs.2014.05.025>
- Hawking, S.: God Created the Integers, pp. 7 (2005). ISBN-13: 978-0-14-101878-2
- Kumar, V.: Quantum Hole (Unpublished)
- Scully, M.O., Englert, B.G., Walther, H.: Quantum optical tests of complementarity. *Nature* **351**, 111–116, 9 May (1991). <https://doi.org/10.1038/351111a0>
- Rosenfeld, L.: Foundations of quantum theory and complementarity. *Nature* **190**(4774), 384–388, 29 April (1961). <https://doi.org/10.1038/190384a0>
- Kumar, V.: On the electro-optic couple systems of quantum station for quantum communication based on phase shift scheme. *Lecture Notes in Electrical Engineering*. Springer (accepted)
- Julsgaard, B., Sherson, J., Cirac, J.I., Fiurasek, J., Polzik, E.S.: Experimental demonstration of quantum memory for light. *Nature* **432**, 482–485, 25 November (2004). <https://www.nature.com/articles/nature03064.pdf>
- Maxwell, J.C.: A dynamical theory of the electromagnetic field. *Phil. Trans. R. Soc. Lond.* **155**, 459–512, 1 January (1865). <http://www.bem.fi/library/1865-001.pdf>
- Hertz, H.: On the Finite Velocity of Propagation of Electromagnetic Actions (Sitzungsber. d. Berl. Akad. d. Wiss. Feb. 2, 1888. Wiedemann's Ann. **34**), English Translation by

- D.E. Jones. <http://onlinebooks.library.upenn.edu/webbin/book/lookupname?key=Hertz%2C%20Heinrich%2C%201857-1894>
- 16. Sadiku, M.N.O.: Principles of Electromagnetics, pp. 377, 4th edn. 21st impression 2015. ISBN-13: 978-0-19-806229-5
 - 17. Qiu, J.: Quantum communications leap out of the lab. *Nature* **508**, 441–442, 24 April (2014). https://www.nature.com/polopoly_fs/1.15093!/menu/main/topColumns/topLeftColumn/pdf/508441a.pdf
 - 18. Hu, J.Y., Yu, B., Jing, M.Y., Xiao, L.T., Jia, S.T., Qin, G.Q., Long, G.L.: Experimental quantum secure direct communication with single photons. *Nature Light Sci. Appl.* 1–5, September (2016). <https://www.nature.com/articles/lsa2016144.pdf>
 - 19. Nielsen, M.A.: Simple rules for a complex quantum world. *Sci. Am.*, 25–33, November (2002). <http://michaelnielsen.org/papers/SciAmSimpleRules.pdf>

Realization of Fractional-Order Proportional Derivative Controller for a Class of Fractional-Order System in Delta Domain



Jaydeep Swarnakar, Prasanta Sarkar and Lairenlakpam Joyprakash Singh

Abstract The study of the fractional-order controller (FOC) is an emerging area of research in system and control at the present scenario. In this paper, a fractional-order proportional derivative (PD^η) controller is implemented for a class of fractional-order plant from certain design specifications of the frequency domain. The entire work has been accomplished in two stages. The first stage deals with the design of continuous-time FOC to satisfy the desired design criterions. In the second stage, a continued fraction expansion (CFE) based direct discrete-time approximation method has been employed to realize the continuous-time FOC in delta domain. The realization of the discrete-time FOC employs the transformed delta operator as a generating function. The efficacy of the controller design method is verified by engaging the discretized FOC with the fractional-order plant and subsequently verifying the frequency response of the discretized open-loop system with respect to the desired frequency response of the original open-loop fractional-order system. The robustness of the overall controlled system is also investigated by altering the plant gain. Simulation results are presented to justify the efficacy of the proposed FOC realization method.

Keywords Fractional-order controller · Continued fraction expansion (CFE) · Delta operator · Delta domain

J. Swarnakar (✉) · L. J. Singh

Department of Electronics & Communication Engineering, School of Technology, North-Eastern Hill University, Shillong 793022, Meghalaya, India
e-mail: jaydeepswarnakar@gmail.com

L. J. Singh

e-mail: jplairen@gmail.com

P. Sarkar

Department of Electrical Engineering, National Institute of Technical Teachers' Training and Research, Kolkata 700106, West Bengal, India
e-mail: psarkar@nittrkol.ac.in

1 Introduction

The study of the fractional-order system has been emerged remarkably since the last few decades among the scientists, engineers or research community as a whole, although the subject fractional calculus was mainly confined as a theoretical field of research among the mathematicians before the nineteenth century [1]. In the field of control engineering, the advantage of employing FOC has been well established through the use of CRONE controller, noninteger-order PID controller leading to improve the performance of the overall controlled system in terms of robustness [2, 3]. FOC is generally employed with the integer-order plant as the plant model is already available in the classical sense [4, 5]. However, the FOC has also been employed with the fractional-order plant for enhancing the system control performance [6–8]. Realization of the infinite-dimensional FOC within a finite band is fundamental from a practical point of view. For continuous-time realization, the FOC is first approximated in a finite band using rational approximation method of S domain. Oustaloup approximation has been widely employed for the said purpose. Digital realization of FOC can be done by using indirect approach and direct approach. For both the approaches, the continuous-time FOC is developed initially from the given design specifications like gain crossover frequency, phase margin, etc. In case of indirect realization, FOC is first converted to the equivalent integer-order transfer function of S domain, and then it is discretized [9]. In case of direct realization, FOC is directly converted to a rational transfer function of conventional z domain. For this purpose, one generating function (e.g. Tustin operator, Al-Alaoui operator) is needed to choose primarily. Then, the generating function is expanded to obtain the discrete-time approximation of FOC as a whole [10, 11].

The discrete-time system is generally modelled using shift operator or its analogous frequency variable z . But, the shift operator-based model fails to maintain the parity between the discrete-time system and its continuous-time counterpart in the proper sense at the lower sampling time limit. Middleton and Goodwin first proposed an alternate modelling based on delta operator for representing the discrete-time system and the continuous-time system under a unified framework [12]. Delta operator uses the principle of signal differencing as oppose to signal shifting and bridges the gap between the discrete-time system and its continuous-time counterpart so that the original continuous-time system can be retrieved from the corresponding discrete-time model when sampling time tends to zero [13].

In this paper, a fractional-order proportional derivative controller is realized in delta domain for a class of fractional-order plant. The direct realization method of FOC is adopted in this work. In the first stage, the continuous-time FOC is developed from the specified frequency domain design criterions. Then, the continuous-time FOC is subsequently realized in delta domain. The realization is based on the CFE approximation method where the delta operator has been used as a generating function. The paper is organized as follows: Sect. 1 deals with the introduction. In Sect. 2, the entire design process of the fractional-order proportional derivative controller along with its discrete-time realization is described in delta domain. In

Sect. 3, the performance of the discretized FOC has been compared with respect to its continuous-time counterpart by involving it with the fractional-order plant model. The robustness of the overall controlled system is also studied in delta domain by varying the plant gain. Relevant simulation results are presented. Section 4 concludes the whole work.

2 Design of Fractional-Order Controller

The fractional-order plant considered over here has been used to model the dynamics of many biological materials as seen from the existing literature [14]. The transfer function of the fractional-order plant is given as follows:

$$P(s) = \frac{K}{s(\tau s^\rho + 1)} \quad (1)$$

A fractional-order proportional derivative controller is employed to control the aforementioned fractional-order plant satisfying the following design specifications:

- Phase margin = φ_m^0
- A flat phase characteristics around the gain crossover frequency to ensure the robustness against the variation of the plant gain.
- Gain crossover frequency = ω_{cg} rad/s.

The transfer function of the fractional-order proportional derivative controller is given as below

$$C(s) = K_p(1 + K_d s^\eta) \quad (2)$$

The open-loop transfer function of the system can be written as follows:

$$H(s) = P(s)C(s) = \{KK_p(1 + K_d s^\eta)\}/\{s(\tau s^\rho + 1)\} \quad (3)$$

As per the aforementioned design criterions, the following relationship can be established accordingly:

$$\text{Arg}\{H(j\omega_{cg})\} = \text{Arg}\{P(j\omega_{cg})C(j\omega_{cg})\} = -\pi + \phi_m \quad (4)$$

$$\left\{ \frac{d\text{Arg}H(j\omega)}{d\omega} \right\}_{\omega=\omega_{cg}} = \left[\frac{d\text{Arg}\{P(j\omega)C(j\omega)\}}{d\omega} \right]_{\omega=\omega_{cg}} = 0 \quad (5)$$

$$|H(j\omega_{cg})| = |P(j\omega_{cg})C(j\omega_{cg})| = |P(j\omega_{cg})||C(j\omega_{cg})| = 1 \quad (6)$$

In [6], it is shown that the design criterions corresponding to Eqs. (4) and (5) cannot be satisfied simultaneously by using classical proportional derivative controller. However, a fractional-order controller can be useful to handle such design criterions for a class of fractional-order plant considered over here. By solving the Eqs. (4), (5) and (6), the values of K_p , K_d and η can be derived and subsequently, the continuous-time FOC can be obtained. The continuous-time fractional-order controller thus obtained from the above design method has been realized in delta domain.

The δ -operator is denoted as follows:

$$\delta = \frac{q - 1}{\Delta} \quad (7)$$

where Δ and q indicate the sampling period and the forward shift operator, respectively. The delta operator unifies to the continuous-time derivative operator at the lower sampling time as shown below

$$\lim_{\Delta \rightarrow 0} \delta \psi(t) = \lim_{\Delta \rightarrow 0} \frac{\psi(t + \Delta) - \psi(t)}{\Delta} = \frac{d}{dt} \psi(t) \quad (8)$$

Similar relationship is observed in the complex domain as well. The frequency variable ‘ γ ’ of delta domain is associated with the discrete-time frequency variable ‘ z ’ as given below

$$\gamma = (z - 1)/\Delta \quad (9)$$

The correspondence between s (continuous-time frequency variable) and γ can be established by replacing $z = e^{s\Delta}$ in Eq. (9) as given below

$$s = \{\ln(1 + \Delta\gamma)\}/\Delta \quad (10)$$

Equation (10) represents the delta operator in the complex domain which indicates s to γ transform. But, this relationship does not reflect a rational function of delta domain. Therefore, the original delta operator has been reconstructed to frame the suitable generating function in the form of a rational function by replacing the exponential term $e^{s\Delta}$ by its first order approximation using trapezoidal rule as shown below

$$\begin{aligned} e^{s\Delta/2}/e^{-s\Delta/2} &= \Delta\gamma + 1 \\ &\Rightarrow (1 + s\Delta/2)/(1 - s\Delta/2) \approx \Delta\gamma + 1 \\ &\Rightarrow s \approx (2\gamma)/(2 + \Delta\gamma) \end{aligned} \quad (11)$$

Equation (11) represents the transformed delta operator which acts as a generating function for approximating the FOC in this work. Using Eq. (11), the fractional-order operator s^η is represented as below

$$s^\eta = \{(2\gamma)/(2 + \Delta\gamma)\}^\eta \quad (0 < \eta < 1) \quad (12)$$

In this paper, the FOC is realized in delta domain by taking the fifth-order CFE approximation of $\{(2\gamma)/(2 + \Delta\gamma)\}^\eta$. The reason behind choosing the fifth-order approximation is to obtain the reasonable magnitude approximation and phase approximation for the fractional power terms and so as to obtain the moderate approximation for the overall controlled system. The mathematical representation of CFE approximation is given as follows [15]:

$$(1 + p)^\eta = \left[1; \frac{\eta p}{1}, \frac{(1 - \eta)p}{2}, \frac{(1 + \eta)p}{3}, \frac{(2 - \eta)p}{2}, \frac{(2 + \eta)p}{5}, \frac{(3 - \eta)p}{2}, \frac{(3 + \eta)p}{7}, \dots \right] \quad (13)$$

The fifth-order discrete-time approximations for all the fractional power terms (s^ρ and s^η) existing in the overall controlled system have been obtained in delta domain replacing p by $\lceil \{(2\gamma)/(2 + \Delta\gamma)\} - 1 \rceil$ in Eq. (13). $P(s)$, $C(s)$ and $H(s)$ are discretized to obtain $P_\delta(\gamma)$, $C_\delta(\gamma)$ and $H_\delta(\gamma)$ in delta domain and finally, the frequency responses of $H(s)$ and $H_\delta(\gamma)$ have been compared to evaluate the efficacy of the proposed FOC realization method in delta domain.

3 Simulation and Results

Taking $K = 1$, $\tau = 0.4$ and $\rho = 1.4$, the transfer function of the fractional-order plant is obtained as below

$$P(s) = \frac{1}{s(0.4s^{1.4} + 1)} \quad (14)$$

Let a fractional-order proportional derivative controller be employed for controlling the fractional-order plant to satisfy the following design specifications:

- Phase margin = 80°
- A flat phase characteristics around the gain crossover frequency.
- Gain crossover frequency = 8 rad/s.

By solving the Eqs. (4), (5) and (6) for the given design specifications, we obtain $K_d = 0.4865$, $\eta = 1.297$ and $K_p = 7.9831$. Therefore, the transfer function of the continuous-time FOC is acquired as follows:

$$C(s) = 7.9831(1 + 0.4865s^{1.297}) \quad (15)$$

Now, the continuous-time FOC is realized in delta domain given by Eq. (16) as below

$$C_\delta(\gamma) = 7.9831 \left[1 + 0.4865 \{2\gamma/(2 + \Delta\gamma)\} \{2\gamma/(2 + \Delta\gamma)\}^{0.297} \right] \quad (16)$$

Table 1 Fifth-order CFE approximation of the fractional power terms in delta domain taking $\Delta = 0.001$ s

Fractional power term	Fifth-order approximation
$\{2\gamma/(2 + \Delta\gamma)\}^{0.297}$	$\frac{7219\gamma^5 + 130100\gamma^4 + 418700\gamma^3 + 343000\gamma^2 + 67950\gamma + 1803}{1837\gamma^5 + 68290\gamma^4 + 343500\gamma^3 + 418500\gamma^2 + 129700\gamma + 7154}$
$\{2\gamma/(2 + \Delta\gamma)\}^{0.4}$	$\frac{8757\gamma^5 + 143100\gamma^4 + 428600\gamma^3 + 327500\gamma^2 + 59520\gamma + 1323}{1353\gamma^5 + 59850\gamma^4 + 328000\gamma^3 + 428400\gamma^2 + 142700\gamma + 8686}$

Using Eqs. (14) and (15), the transfer function of the open-loop controlled system is obtained as below

$$H(s) = P(s)C(s) = \frac{7.9831(1 + 0.4865s^{1.297})}{s(0.4s^{1.4} + 1)} \quad (17)$$

Employing the discrete-time FOC shown in Eq. (16) with the discretized fractional-order plant model, the following open-loop transfer function is obtained in complex delta domain.

$$H_\delta(\gamma) = P_\delta(\gamma)C_\delta(\gamma) = \frac{7.9831[1 + 0.4865\{2\gamma/(2 + \Delta\gamma)\}\{2\gamma/(2 + \Delta\gamma)\}^{0.297}]}{\{2\gamma/(2 + \Delta\gamma)\}[0.4\{2\gamma/(2 + \Delta\gamma)\}\{2\gamma/(2 + \Delta\gamma)\}^{0.4} + 1]} \quad (18)$$

Using Table 1 and taking $\Delta = 0.001$ s, Eq. (18) is approximated as follows:

$$H_\delta(\gamma) \approx \frac{75890\gamma^{12} + 1.565 \times 10^8\gamma^{11} + 9.616 \times 10^9\gamma^{10} + 1.737 \times 10^{11}\gamma^9 + 1.275 \times 10^{12}\gamma^8 + 4.758 \times 10^{12}\gamma^7 + 1.024 \times 10^{13}\gamma^6 + 1.323 \times 10^{13}\gamma^5 + 9.964 \times 10^{12}\gamma^4 + 4.061 \times 10^{12}\gamma^3 + 8.182 \times 10^{11}\gamma^2 + 6.882 \times 10^{10}\gamma + 1.984 \times 10^9}{\gamma \left(\begin{array}{l} 2.574 \times 10^7\gamma^{11} + 1.388 \times 10^9\gamma^{10} + 2.253 \times 10^{10}\gamma^9 + 1.53 \times 10^{11}\gamma^8 \\ + 5.467 \times 10^{11}\gamma^7 + 1.174 \times 10^{12}\gamma^6 + 1.551 \times 10^{12}\gamma^5 + 1.201 \times 10^{12}\gamma^4 \\ + 4.995 \times 10^{11}\gamma^3 + 1.018 \times 10^{11}\gamma^2 + 8.605 \times 10^9\gamma + 2.486 \times 10^8 \end{array} \right)} \quad (19)$$

The frequency responses of $C(s)$ and $C_\delta(\gamma)$ are shown in Fig. 1. The frequency responses of the continuous-time FOC and the corresponding discrete-time FOC are very similar in nature. The frequency responses of $H(s)$ and $H_\delta(\gamma)$ are shown together in Fig. 2. From Fig. 2, we obtain $\omega_{cg} = 8$ rad/s, $\varphi_m = -100^\circ + 180^\circ = 80^\circ$ and the phase curve seems to be flat around gain crossover frequency. So, all the prerequisite design specifications are satisfied. It is also apparent that the magnitude response and the phase response of $H_\delta(\gamma)$ are almost similar to the frequency response of $H(s)$. Therefore, the delta operator based discrete-time realization of FOC behaves almost as a continuous-time FOC and thereby meeting all the specified design criterions when it is connected to the fractional-order plant. The frequency responses of the open-loop system $H_\delta(\gamma)$ with FOC $C_\delta(\gamma)$ have been tested

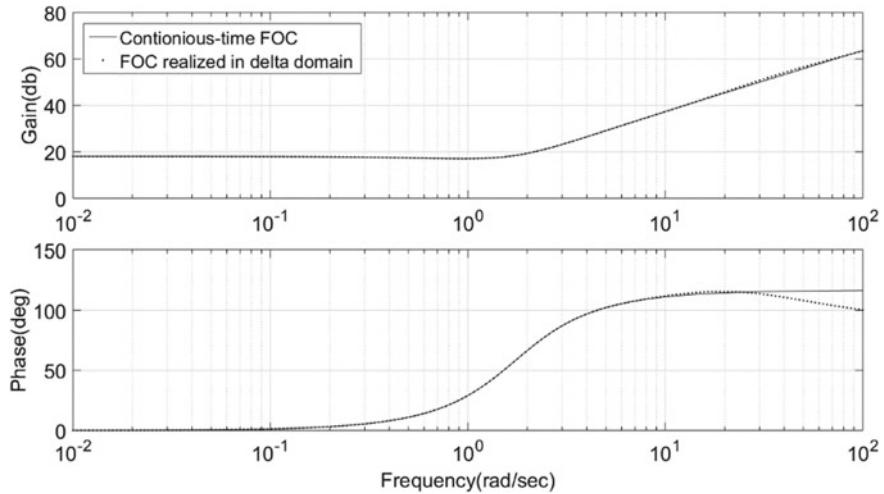


Fig. 1 Comparison between the frequency responses of $C(s)$ and $C_\delta(\gamma)$

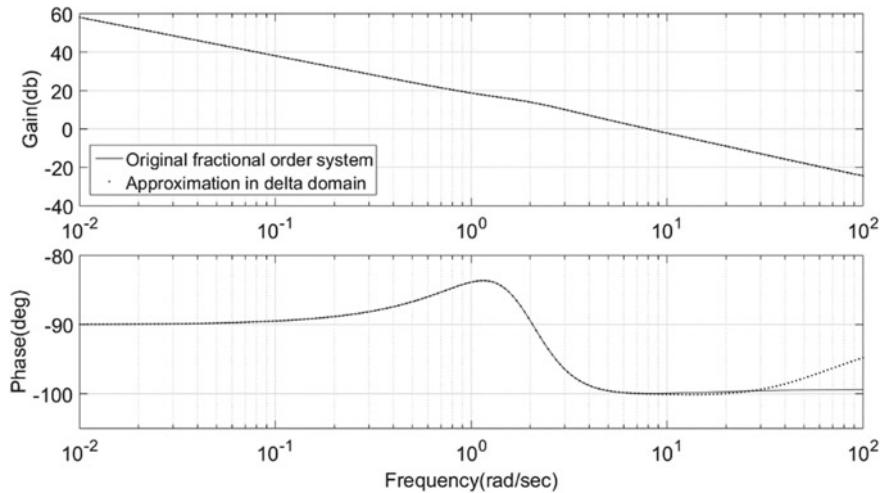


Fig. 2 Comparisons between the frequency responses of $H(s)$ and $H_\delta(\gamma)$ with FOCs of respective domains

by increasing the original plant gain K up to 200% as shown in Fig. 3. It has been noticed that the phase margin does not change even though the plant gain is varied. Therefore, the FOC realized in delta domain offers robust characteristics to the closed loop system against the variations of uncertain plant gain as apparent from the frequency response characteristics. Another important aspect regarding the FOC realization in delta domain is that the frequency response of the discrete-time fractional-order system exactly corresponds to the frequency response of the origi-

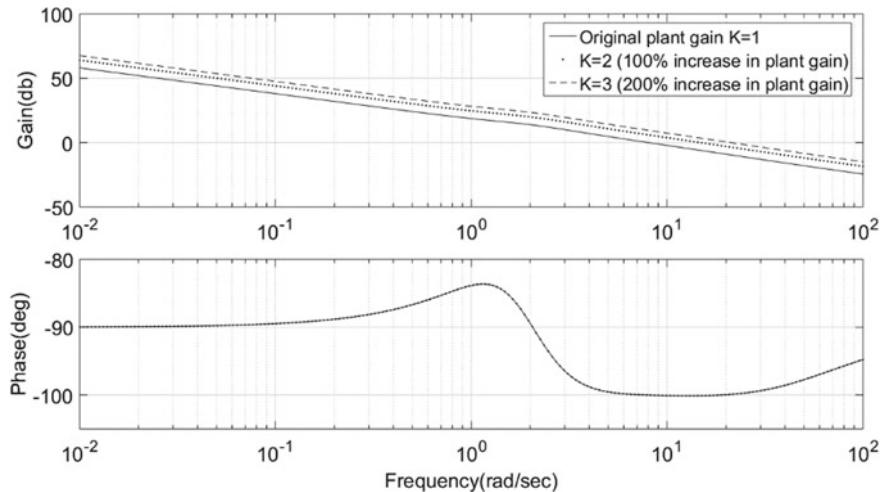


Fig. 3 Frequency responses of $H_\delta(\gamma)$ with FOC $C_\delta(\gamma)$ for different values of plant gain

nal continuous-time fractional system at the lower sampling time $\Delta = 0.001$ s. This clearly signifies the unification of both the analog and digital fractional-order system together at a high sampling frequency.

4 Conclusions

In this paper, a discrete-time approximation method is presented to realize a fractional-order proportional derivative controller for a class of fractional-order system. At first, the continuous-time FOC is developed from some given design specifications and later on the FOC is realized in delta domain by using the CFE approximation method. Simulation results establish that the FOC realized using the delta operator exactly corresponds to the continuous-time FOC and therefore, all the desired control objectives like gain crossover frequency, phase margin, flatness of the phase response characteristics around the gain crossover frequency are fulfilled by the overall discretized controlled system as met by its continuous-time counterpart. The robustness of the overall controlled system is also investigated in delta domain by varying the plant gain. It has also been perceived that the frequency response of the discretized controlled system of delta domain unifies to the frequency response of the original continuous-time system at very low sampling time limit. This justifies the usefulness of the proposed method for realizing FOC in delta domain instead of conventional z domain to meet the specified design criterions as obtained by using continuous-time FOC.

References

1. Monje, C.A., Chen, Y.Q., Vinagre, B.M., Xue, D., Feliu, V.: Fractional-Order Systems and Controls. Springer, London (2010)
2. Oustaloup, A.: CRONE Control: Robust Control of Non-integer order. Hermès, Paris (1991)
3. Podlubny, I.: Fractional-order systems and $PI^{\lambda}D^{\mu}$ controllers. IEEE Trans. Autom. Control **44**(1), 208–214 (1999)
4. Folea, S., De Keyser, R., Birs, I.R., Muresan, C.I., Ionescu, C.: Discrete-time implementation and experimental validation of a fractional order PD controller for vibration suppression in airplane wings. Acta Polytechnica Hungarica **14**(1), 191–206 (2017)
5. Sun, J., Wang, C., Xin, R.: Design of fractional order proportional differentiation controller for second order position servo system. In: Proceedings of the IEEE Chinese Control and Decision Conference (CCDC), pp. 5939–5944 (2018)
6. Luo, Y., Li, H., Chen, Y.: Fractional order proportional and derivative controller synthesis for a class of fractional order systems: tuning rule and hardware-in-the-loop experiment. In: Proceedings of the 48th IEEE Conference on Decision and Control Held Jointly with 28th Chinese Control Conference, pp. 5460–5465 (2009)
7. Luo, Y., Chen, Y.Q., Wang, C.Y., Pi, Y.G.: Tuning fractional order proportional integral controllers for fractional order systems. J. Process Control **20**(7), 823–831 (2010)
8. Wang, C., Fu, W., Shi, Y.: Tuning fractional order proportional integral differentiation controller for fractional order system. In: Proceedings of the 32nd IEEE Chinese Control Conference (CCC), pp. 552–555 (2013)
9. Zamojski, M.: Implementation of fractional order PID controller based on recursive oustaloup'e filter. In: Proceedings of the IEEE International Interdisciplinary Ph.D. Workshop, pp. 414–417 (2018)
10. Chen, Y.Q., Vinagre, B.M., Podlubny, I.: Continued fraction expansion approach to discretizing fractional order derivatives—an expository review. Nonlinear Dyn. **38**(1–4), 155–170 (2004)
11. Song, B., Xu, L., Lu, X.: A comparative study on Tustin rule based discretization methods for fractional-order differentiator. In: Proceedings of the 4th IEEE International Conference on Information science and Technology, Shenzhen, China, pp. 515–518 (2014)
12. Middleton, R.H., Goodwin, G.C.: Digital Control and Estimation: A Unified Approach. Prentice Hall, Englewood Cliffs, NJ (1990)
13. Swarnakar, J., Sarkar, P., Singh, L.J.: Realization of fractional-order operator in complex domains—a comparative study. In: Bera, R., Sarkar, S., Chakraborty, S. (eds.) Lecture Notes in Electrical Engineering, vol. 462, pp. 711–718. Springer, Singapore (2018)
14. Magin, R.L.: Fractional Calculus in Bioengineering. Begell House Publishers Inc., Redding (2006)
15. Swarnakar, J., Sarkar, P., Singh, L.J.: A unified direct approach for discretizing fractional-order differentiator in delta-domain. Int. J. Modeling, Simul. Sci. Comput. **9**(6), 185055-1–185055-20 (2018)

Modelling of Microwave-Based Regeneration in Composite Regeneration Emission Control System



Caneon Kurien, Ajay Kumar Srivastava, Karan Anand and Niranjan Gandigudi

Abstract Fuel-based regeneration of the particulate filter channels in post-treatment emission control system is leading to uncontrolled combustion causing damage to filter substrate. One of the major challenges faced by the automotive industry is the development of an alternate technique for regeneration of accumulated soot particles. Electromagnetic radiations in microwave region have been proposed in this paper as a suitable alternative for regeneration of trapped soot particles. A detailed literature study has been conducted on the heating principle and governing equations for microwave radiations. Modelling and simulation of microwave-based emission control system have been conducted using COMSOL Multiphysics. A comparative study has been conducted for two models with different positions of waveguide for the introduction of microwave radiations. The results of simulation for the model with waveguide at centre position showed that there is an even distribution of electric field in the system and the cordierite-based filter substrate is suitable for microwave-based regeneration system owing to its dielectric properties.

Keywords Emission · Filter · Microwave · Regeneration · Soot

C. Kurien (✉) · A. K. Srivastava

Department of Mechanical Engineering, University of Petroleum and Energy Studies, Dehradun 248007, India

e-mail: ckurien@ddn.upes.ac.in

A. K. Srivastava

e-mail: akumar@ddn.upes.ac.in

K. Anand · N. Gandigudi

Department of Aerospace Engineering, University of Petroleum and Energy Studies, Dehradun 248007, India

e-mail: kanand.mech94@gmail.com

N. Gandigudi

e-mail: gandigudiniru@gmail.com

1 Introduction

Extensive use of fossil fuel-based vehicles in the automobile sector has increased the concern over environment, due to the adverse impact of its exhaust emissions on environment and human health [1]. Emissions produced by internal combustion engines include carbon dioxide, carbon monoxide, nitrates, unburnt hydrocarbons, sulphur oxide and particulate matter [2]. Exposure to carbon monoxide gas will lead to dizziness, vomiting tendency and also blocks haemoglobin in blood by forming carboxyhaemoglobin [3]. The European Union Research Organization (EURO) has imposed stringent regulations on the maximum permissible limits of exhaust emissions from the internal combustion engines [4, 5]. The EURO I standard was introduced in 1994 and with time, the limits were reduced in a step-by-step phenomenon and presently, EURO VI standard is introduced, which has reduced the overall emissions by more than 80% as compared to that of EURO I standard [6]. The toxic gases like carbon monoxide and hydrocarbons are produced due to incomplete combustion of fuel in the engine cylinder [7]. Complete combustion of the fuel is not possible since it is difficult to achieve thermodynamic equilibrium due to various reasons like limited time for chemical oxidation, lack of homogeneity in mixture and lack of uniformity in temperature distribution [8]. Presence of sulphur content in the fuel is the sole reason for the emission of sulphur oxide [9]. Various modifications have to be done on the present-day internal combustion engines to reduce its emissions to meet the stringent emission norms [10]. Some of these modifications include cylinder alterations, injection strategies, exhaust gas recirculation, use of alternate fuels and post-treatment emission control systems [11]. Post-treatment emission control system includes diesel oxidation catalysis and diesel particulate filtration systems [12]. Oxidation of carbon monoxide and reduction of hydrocarbons takes place in the oxidation catalysis system [13].

Particulate filtration system traps the particulate matter along the walls of its alternately plugged filter channels. The accumulated particles have to be regenerated to avoid rise in back pressure and clogging of filter channels [14]. In commercially available particulate filtration systems, regeneration of accumulated particles are conducted by injecting fuel to the filter channels which will self-ignite by the temperature of exhaust gas and burns the trapped particles [15]. But the fuel-based regeneration will lead to uncontrolled combustion inside the filter channels and eventually damage the filter substrate reducing its service life [16]. Development of alternate regeneration technique is one of the major challenges faced by the automotive industry and microwave radiations-based regeneration have been proposed in this paper to be a suitable alternative technique. Microwave radiations have gained its popularity in food processing and chemical industry owing to its fast and volumetric heating properties [17]. The advantage of microwave radiations is that it heats the target material by volumetric heating, whereas the conventional technique does it by surface heating [18]. Distribution of temperature while microwave heating can be predicted by the development of mathematical model using electromagnetic and heat transfer equations. One-dimensional analysis on power decay of radiations emitted by magnetron

can be determined by analytical equations of Lambert's law [19]. Microwave radiations also don't require any medium for transferring energy from source to target material.

2 Microwave Heating Principle

The major components of electromagnetic waves include electric field and magnetic field. Electric and magnetic field changes its position in space when a corresponding electric or magnetic charge in the medium changes its position [20]. The changes in electric and magnetic fields will result in the production of oscillatory waves named as electromagnetic waves. The properties of electromagnetic waves are determined by its frequency, field strength, speed of light and wavelength [21]. Electromagnetic radiations were discovered in 1873 by Maxwell and the orientation of elements was established by right-hand thumb rule. The properties of electromagnetic waves are related by Eq. 1, where c = speed of light (m/s); f = frequency (Hz) and λ = wavelength.

$$c = f\lambda \quad (1)$$

Polarization of molecules in the frequency range between 300 MHz and 300 GHz will result in heating of dielectric material. Penetration of microwave energy to the target material will result in molecular friction resulting in dipolar rotation of the molecules producing volumetric distribution of heat. Major parameters determining the interaction of electric field and material are dielectric constant and dielectric loss factor as shown in Eq. 2, where dielectric constant determines the ability of material to store electric energy and dielectric loss factor determines the dissipation of electric energy to heat. The materials are classified into low loss and medium loss materials based on its dielectric loss tangent, which is determined by Eq. 3. The electrical conductivity is given by Eq. 4, where f is frequency (Hz) and ϵ_0 is the free space permittivity ($8.54 * 10^{-12}$ F/m) [17].

$$\epsilon^* = \epsilon' - j\epsilon'' \quad (2)$$

$$\tan \delta = \frac{\epsilon''}{\epsilon'} \quad (3)$$

$$\sigma = 2\pi f \epsilon_0 \epsilon'' \quad (4)$$

Power per unit volume (W/m^3) required for conversion of the electric properties of microwave radiation for dissipation of power in a lossy material is given by Eq. 5, where E is the electric field strength of the target material and f is the frequency in GHz [22].

$$P_v = 2\pi f \epsilon_0 \epsilon'' E^2 \quad (5)$$

One-dimensional microwave power dissipation inside the target material can be calculated using Lambert's law, which is a simplified approach without considering the electric field distribution. It assumes that the radiations incident on the material are normal to the surface and energy dissipated by the radiations exponentially as stated in Eq. 7, where power flux at surface (W/m^2), d_p is the penetration depth and z is the spatial distance in z direction. The depth of microwave penetration is dependent on the dielectric properties of the material as shown in Eq. 8 [21].

$$F = F_o \exp\left(\frac{-z}{d_p}\right) \quad (6)$$

$$d_p = \frac{C_o}{2\pi f} [2 \epsilon' ((1 + (\frac{\epsilon''}{\epsilon'})^2)^{0.5} - 1)]^{-0.5} \quad (7)$$

The microwave heating of material by considering variations in electric and magnetic fields are governed by Maxwell equations. The relation between electric field intensity and magnetic induction is given by Faraday's law as shown in Eq. 8, where E is electric field intensity, B is magnetic induction, H is magnetic field intensity, ω is angular frequency, μ_m is permeability and ϵ is the permittivity. The second governing equation is to determine the magnetic field intensity which is the sum of current flux and rate of change of electric displacement as shown in Eq. 9, where J is current flux, The other two governing equations for finding the electric displacement (D) and magnetic induction (B) is as shown in Eq. 10 [23].

$$\nabla E = -\frac{\partial B}{\partial t} = -\frac{\partial(\mu_m \omega H(t))}{\partial t} \quad (8)$$

$$\nabla H = J + \frac{\partial D}{\partial t} = \sigma \omega E(t) + \frac{\partial(e \omega E(t))}{\partial t} \quad (9)$$

$$\nabla D = q \text{ and } \nabla B = 0 \quad (10)$$

3 Modelling and Analysis

The heat transfer pattern during regeneration of diesel particulate filtration system has been studied in this article using CFD software COMSOL Multiphysics [23]. The filter substrate of the diesel particulate filtration system is made of cordierite-based ceramic material and its associated properties have been recorded in Table 1. The filter substrate is contained in a metallic box made of GI sheet, which acts as an oven cavity and whose dimensions are $26.7 \text{ cm} \times 27 \text{ cm} \times 18.8 \text{ cm}$. Microwave radiations are introduced into the oven cavity by placing a magnetron port in one of the faces of the box.

Table 1 Properties of cordierite substrate used for diesel particulate filtration of exhaust gas

Properties	Value	Unit
Relative permeability	1	–
Relative permittivity	$2.9 + 0.14j$	–
Electrical conductivity	0	S/m
Thermal conductivity	3.0	W/m K
Density	2600	kg/m ³
Heat capacity	1465.38	J/kg K

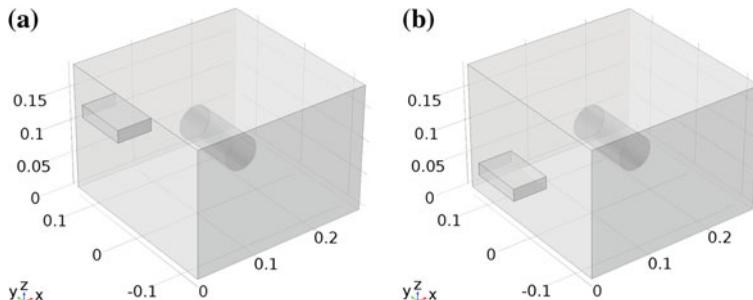


Fig. 1 **a** Model of microwave-based emission control system with magnetron port at top. **b** Model of microwave-based emission control system with magnetron port at centre

The frequency of the excitation wave is 2.45 GHz and the port input power is 1 kW. The prime objective of this simulation study is to determine the pattern of electric field distribution and temperature distribution in the oven cavity with cordierite filter substrate at its centre. Magnetron port is assumed to be a rectangular waveguide operating at the TE10 mode. The initial temperature of the filter is assumed to be 50 °C and the simulation is carried out for a time period of 120 s post-excitation. For studying the microwave excitation, the frequency domain module of electromagnetic waves are selected with steady-state conditions. Module for heat transfer in solids (h_t) is selected in Time-Dependent mode to study the temperature rise in the DPF.

Two models are considered in this study for determining the temperature distribution and electric field distribution inside the oven cavity of microwave-based regeneration system. One with the magnetron port located at the top of the oven cavity and the other with the port located at the centre of the wall as shown in Fig. 1a, b, respectively. The study solves for the electromagnetic field in the box and the corresponding temperature by coupling the Maxwell's equations and the general energy transport equation for temperature. The mesh for solving the numerical equations has been generated automatically using one-fifth of the wavelength for the electromagnetic wave as the maximum mesh element control parameter.

The contours of electric field distribution after the steady-state simulation of microwave excitation from the magnetron for both cases with the port located at top and centre are as shown in Fig. 2a, b. It can be inferred that the distribution

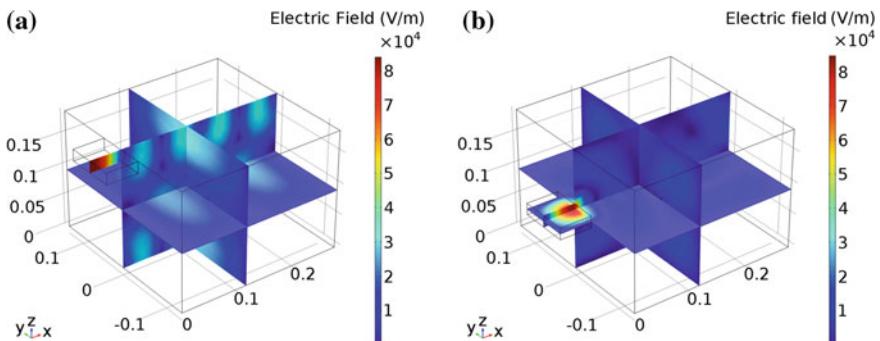


Fig. 2 **a** Contour of electric field distribution in different planes for microwave-based emission control system with magnetron port at top. **b** Contour of electric field distribution in different planes for microwave-based emission control system with magnetron port at centre

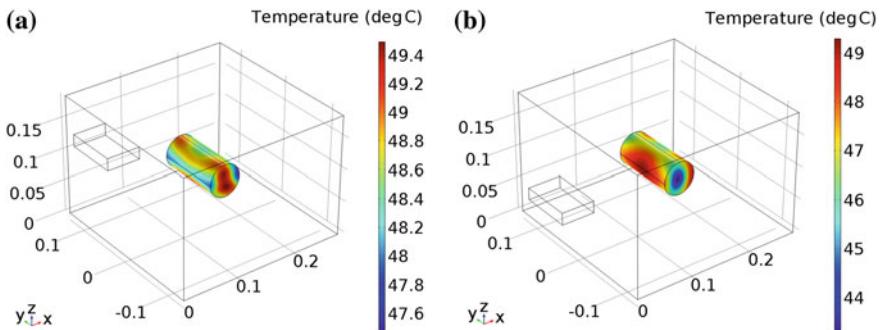


Fig. 3 **a** Contour of temperature distribution in filter substrate of microwave-based emission control system with magnetron port at top. **b** Contour of temperature distribution in filter substrate of microwave-based emission control system with magnetron port at centre

of electric field is more uniform in the case where the port is located at the centre. The uniformity of electric field makes the second case, the one waveguide port at centre as a more effective alternative for microwave-based regeneration system. The temperature distribution for both the cases mentioned above can be observed in Fig. 3a, b. It can be seen that the temperature distribution pattern (owing to difference in electric field distribution) for the both the cases are different, but there are only negligible changes in the magnitude of temperature for both cases. This can be attributed to the fact that cordierite has low sensitivity towards microwave excitation. Also the temperature drop in 120 s is relatively low due to the high heat capacity of cordierite. Hence, with these observations in mind, it can be concluded that ceramic

cordierite-based filter substrate is the most suitable material for the development of emission control system with microwave-based regeneration since it is transparent to microwave radiations.

4 Conclusions

Reducing the toxicity levels of exhaust emissions released by compression ignition engines is one of the major challenges faced by automotive and other industries. Post-treatment emission control system is considered to be the best possible solution for reducing the exhaust emissions to meet the latest emission norms. Fuel-based regeneration of the accumulated soot particles in the walls of diesel particulate filter substrate is affecting its service life and causing early damage due to overheating of the filter substrate. Development of a system, which will burn only the accumulated soot particles without causing any damage to the filter substrate will be the effective alternative for fuel-based regeneration system. Theoretical study on the effectiveness of the microwave-based regeneration system was conducted in this article by modelling and conducting simulations by creating a virtual environment in CFD software COMSOL Multiphysics. Comparative study on the electric field distribution in the system with magnetron port at top and centre was conducted to determine the effective position for uniform distribution of the electric field inside the oven cavity. The contours of electric field distribution showed that the electric field distribution is uniform in the case of the model with magnetron port at the centre. The results of simulation showed that the ceramic cordierite-based filter substrate is transparent to the microwave radiations, since there was no appreciable rise in the temperature of the filter substrate. Hence the theoretical validation of the proposed microwave-based regeneration system is done. The future scope of this work includes simulation of system with accumulated soot particles along the filter wall, so that the depth of penetration and regeneration time can be determined.

References

1. Kang, W., Choi, B., Jung, S., Park, S.: PM and NO_x reduction characteristics of LNT/DPF+SCR/DPF hybrid system. *Energy* **143**, (Supplement C), 439–447 (2018)
2. Kurien, C., Srivastava, A.K.: Active regeneration of diesel particulate filter using microwave energy for exhaust emission control. In: Intelligent Communication, Control and Devices, pp. 1233–1241 (2018)
3. Carder, D., Ryskamp, R., Besch, M., Thiruvengadam, A.: Emissions control challenges for compression ignition engines. *Procedia IUTAM* **20**(X), 103–111 (2017)
4. Suarez-Bertoa, R., Astorga, C.: Impact of cold temperature on Euro 6 passenger car emissions. *Environ. Pollut.* **234**(Supplement C), 318–329 (2018)
5. Kouchachvili, L., Yaici, W., Entchev, E.: Hybrid battery/supercapacitor energy storage system for the electric vehicles. *J. Power Sources* **374**, 237–248 (2018)

6. Kurien, C., Srivastava, A.K.: Advancement in the design of automotive catalytic filter for meeting environmental emission norms key words. *Nat. Environ. Pollut. Technol.* no. 2014 (2018)
7. Seher, S.I., Ess, M.N., Bladt, H., Niessner, R., Eigenberger, G., Nieken, U.: A comparison of diesel soot oxidation rates measured with two different isothermal set-ups. *J. Aerosol Sci.* **91**(x), 94–100 (2016)
8. Bai, S., Chen, G., Sun, Q., Wang, G., Li, G.: Influence of active control strategies on exhaust thermal management for diesel particulate filter active regeneration. *Appl. Therm. Eng.* **119**, 297–303 (2017)
9. Guan, C., Li, X., Liao, B., Huang, Z.: Effects of fuel injection strategies on emissions characteristics of a diesel engine equipped with a particle oxidation catalyst (POC). *J. Environ. Chem. Eng.* **4**(4), 4822–4829 (2016)
10. Kurien, C., Srivastava, A.K., Naudin, J.: Modelling of regeneration and filtration mechanism in diesel particulate filter for development of composite regeneration emission control system. *Arch. Mech. Eng.* **65**(2), 277–290 (2018)
11. Kurre, S.K., Garg, R., Pandey, S.: A review of biofuel generated contamination, engine oil degradation and engine wear. *Biofuels* **8**(2), 273–280 (2017)
12. Lee, Y.K., Kim Y.J., Kim, H.J., Han, B., Hong, W.S., Shin, W.H., Cho, G.B.: Development of electrostatic diesel particulate filtration systems combined with a metallic flow through filter and electrostatic methods. *Int. J. Automot. Technol.* **11**(4), 447–453 (2010)
13. Wilklow-Marnell, M., Jones, W.D.: Catalytic oxidation of carbon monoxide by α -alumina supported 3 nm cerium dioxide nanoparticles. *Mol. Catal.* **439**, 9–14 (2017)
14. Herreros, J.M., Gill, S.S., Lefort, I., Tsolakis, A., Millington, P., Moss, E.: Enhancing the low temperature oxidation performance over a Pt and a Pt-Pd diesel oxidation catalyst. *Appl. Catal. B Environ.* **147**, 835–841 (2014)
15. Ning, J., Yan, F.: Composite control of DOC-out temperature for dpf regeneration. *IFAC-PapersOnLine* **49**(11), 20–27 (2016)
16. Bermúdez, V., Serrano, J.R., Piqueras, P., García-Afonso, O.: Pre-DPF water injection technique for pressure drop control in loaded wall-flow diesel particulate filters. *Appl. Energy* **140**, 234–245 (2015)
17. Halim, S.A., Swithenbank, J.: Simulation study of parameters influencing microwave heating of biomass. *J. Energy Inst.* (2018)
18. Birla, S.L., Pitchai, K.: *Simulation of Microwave Processes*, 2nd edn. Elsevier Ltd. (2017)
19. Falciglia, P.P., Roccaro, P., Bonanno, L., De Guidi, G., Vagliasindi, F.G.A., Romano, S.: A review on the microwave heating as a sustainable technique for environmental remediation/detoxification applications. *Renew. Sustain. Energy Rev.* **95**, 147–170 (2018)
20. Punathil, L., Basak, T.: Microwave heating of food materials: theoretical investigations. In: *Reference Module in Food Science*. Elsevier (2017)
21. Regier, M.: Microwave heating. In: *Reference Module in Food Science*. Elsevier (2017)
22. Curet, S., Rouaud, O., Boillereaux, L.: Heat transfer models for microwave thawing applications. In: *Excerpt from the Proceedings of the COMSOL Users Conference, Paris* (2006)
23. Imenokhoyev, I., Matthes, A., Gutte, H., Walter, G.: Numerical 3D-FEM-simulation made by COMSOL multiphysics of a microwave assisted cleaning system for a diesel sooty particle filter and its experimental validation. In: *Proceeding Book, International COMSOL Multiphysics Conference, Ludwigsburg*, vol. 26, no. 28.10 (2011)

Image Analysis of SEM Micrograph of Co-doped ZnO-Based Oxide Semiconductors



Rana Mukherji, Vishal Mathur and Manishita Mukherji

Abstract Digital image processing deals with the processing of digital images using its standard methods like restoration, noise reduction, segmentation and edge detection. The proposed work represents a novel image processing technique access and validates the findings of scanning electron microscopy micrographs of Co-doped ZnO dilute magnetic semiconductors. Median and Sobel filtering techniques are used to improve the quality of the image whereas the watershed segmentation algorithm is used to determine to approximate particle size as well as distribution.

Keywords Digital image processing · Scanning electron microscopy · Median filter · Sobel filter · Watershed segmentation algorithm

1 Introduction

Scanning Electron microscope (SEM) is among the utmost influential instruments accessible in the field of structural characterization of motley of materials at nanolevel [1, 2]. It is basically an electron microscope, which takes images of the specimen's surface via scanning in a raster scan pattern with high-energy electron beam. The SEM has multitudinous benefits on conventional microscopes. Usually, the micrographs obtained from SEM are assessed manually, moreover, the micrographs are exposed to the variability of issues that may damage the image quality. Therefore, there is a need to develop novel algorithms, which support the qualitative and quantitative interpretation of micrographs. Digital Image Processing (DIP) has provided the appreciable path to acquire the quantitative understanding of the morphological study of micrographs. The DIP segmentation algorithms focus on the inequity amid

R. Mukherji (✉) · V. Mathur
The ICFAI University, Jaipur, India
e-mail: rana.mukherji@jpr.amity

M. Mukherji
Amity University Jaipur, Jaipur, Rajasthan, India

diverse molecular segments along with background wherever image enhancement is helpful for decoding eminence feature in SEM images [3, 4].

Therefore, in this paper, an attempt has been made to analyse an SEM image of Co-doped ZnO dilute magnetic semiconductor. The proposed algorithm includes median and Sobel filtering which improve the quality of the image by reducing the noise level of the image. The watershed segmentation technique is used to assess the homogeneity in partial size as well as distribution.

2 Proposed Methodology and Results

The proposed image processing methodology is used to assess grayscale SEM micrographs which are typified by 2D arrays of grayscale intensity values, with every single value connecting to one pixel in the image. Digital Image Processing (DIP) entails mutation of the intensity values amid 0 (black) and 255 (white) according to mathematical operations for 8-bit images. Various DIP operations include filtering process in which a tiny matrix or kernel is convolved with the image matrix (3×3). The processed image will always be contingent on filter type used. Present work focused on image analysis of Co-doped ZnO SEM micrograph. First, the SEM image is converted into a binary image and then introduced to the median filter. At every single pixel in the image, the filter substitutes the intensity value with the median of all those values in its 3×3 pixel neighbourhood which abolishes isolated noise specks of the image by preserving the intelligibility of edges.

Afterwards, the segmentation process is performed using the watershed algorithm. Though, the SEM image segmentation is a tough assignment, specifically when it has overlapping objects and there is a necessity to determine their shape and size of constituent molecules efficiently. In this regard, the watershed algorithm can be used that provides efficient results with particular characteristics [5]. In accordance to the proposed algorithm, the binary image of the micrograph is converted distance transform image. In distance transformation, the brightness of a pixel is controlled via distance between that molecular point and the nearest pixel of molecular periphery. The city block distance transform is used to compute the said distance. In the subsequent step, topological surface plot of the image is created through distance map observations. In the surface plot, the portions illuminated with brighter pixels represents deeper parts of the molecular segments. Finally, a number of basins are formed thereupon for each of the enchain molecules in the image. All pixels of the watershed ridgeline are detected by completing the said process. In the end, the watershed ridgeline differentiates the molecules [6].

The image analysis outcomes of micrographs of Co-doped ZnO are illustrated in Fig. 1a–d in which interspacing basins are categorised with random colour for better distinction. Since the number of distinct colours are very less, it can be deduced that the agglomerations of specimens unveil homogeneous silhouettes. Figure 2 represents a graphical representation of the particle distribution of the said SEM image. The parameters obtained from the process are illustrated in Table 1.

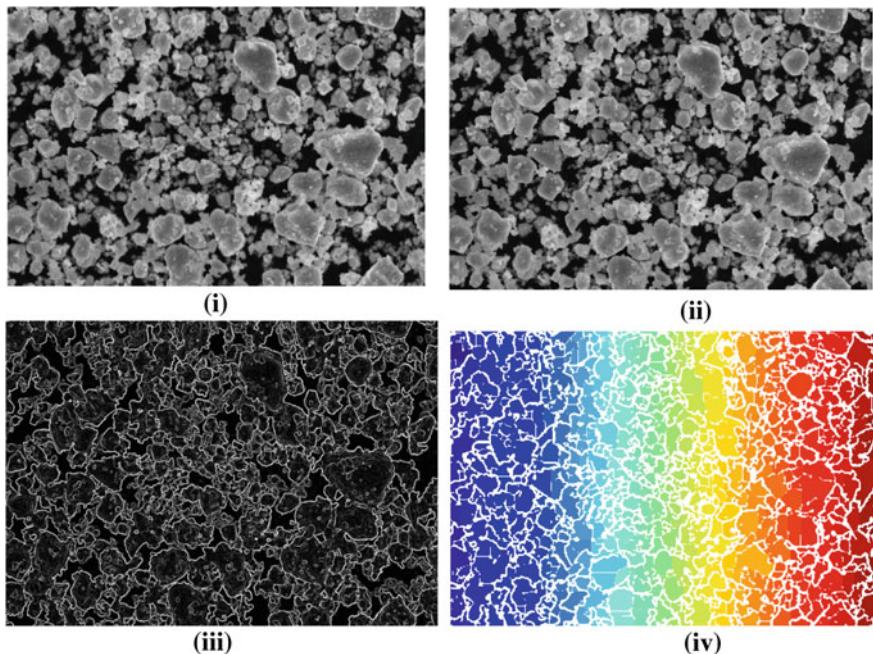


Fig. 1 **a** Original SEM image of co-doped ZnO specimen. **b** Image after the application of median filter. **c** Image after the application of Sobel edge detection filter. **d** Image after the application of watershed algorithm

Fig. 2 Graphical representation of particle distribution of Co-doped ZnO specimens based on watershed segmentation algorithm

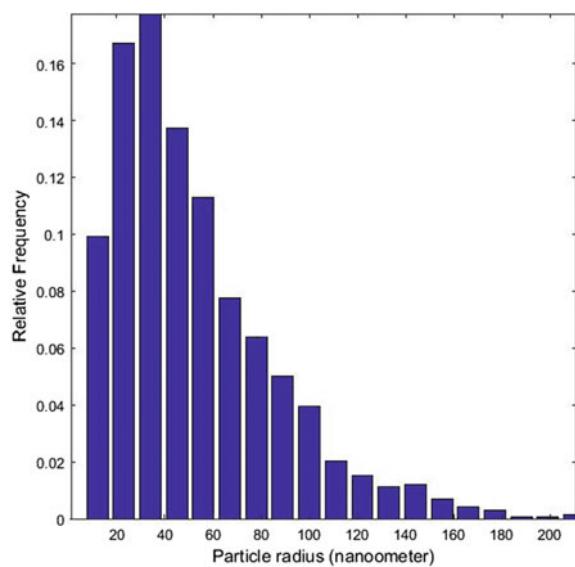


Table 1 Particle-size parameters obtained from analysing 12 random images of each real rock sample calculated by using the three different image processing algorithms

Parameters	Corresponding values
Resolution (nm/pixel)	251.6
Average 2D porosity	20
Average of particle size (nm)	52.20
Average standard deviation of particle size (nm)	33.49

3 Conclusions

The proposed work suggests a novel approach to evaluate the homogeneity particle distribution through SEM images of Co-doped ZnO-based oxide semiconductor specimen. According to the results obtained from analysis, the distribution of particles is homogeneous as suggested by the watershed algorithm result. The particle size varies between 20 and 200 nm, whereas the average size comes out to be 52.20 nm. The average standard deviation of particle size (nm) is found as 33 nm with a resolution of 251.6 nm/pixel.

References

1. Lee, J.H., Yoo, S.I.: An effective image segmentation technique for the SEM image. In: Industrial Technology, 2008. ICIT 2008. IEEE International Conference, pp. 1–5. IEEE Press, New York (2008)
2. Shirazi, S.H., Haq, N., Hayat, K., Naz, S., Haque, I.: Curvelet based offline analysis of SEM images. PLoS ONE **9**(8), e103942 (2014)
3. Barraud, J.: The use of watershed segmentation and GIS software for textural analysis of thin sections. J. Volcanol. Geotherm. Res. **154**(1–2), 17–33 (2006)
4. Galloway, J.A., Montminy, M.D., Macosko, C.W.: Image analysis for interfacial area and cocontinuity detection in polymer blends. Polymer **43**(17), 4715–4722 (2002)
5. Saladra, D., Kopernik, M.: Qualitative and quantitative interpretation of SEM image using digital image processing. J. Microsc. **264**(1), 102–124 (2016)
6. Rabbani, A., Ayatollahi, S.: Comparing three image processing algorithms to estimate the grain-size distribution of porous rocks from binary 2D images and sensitivity analysis of the grain overlapping degree. Spec. Top. Rev. Porous Media Int. J. **6**(1), 71–89 (2015)
7. Rabbani, A., Assadi, A., Kharrat, R., Dashti, N., Ayatollahi, S.: Estimation of carbonates permeability using pore network parameters extracted from thin section images and comparison with experimental data. J. Nat. Gas Sci. Eng. **42**, 85–98 (2017)

The Selection of Options for Closed-Die Forging of Complex Parts Using Computer Simulation by the Criteria of Material Savings and Minimum Forging Force



Volodymyr Kukhar, Elena Balalayeva, Svitlana Hurkovska, Yurii Sahirov, Oleg Markov, Andrii Prysiaznyi and Oleksandr Anishchenko

Abstract The analysis of the technological modes, technical and economic indices of three variants of closed-die forging of the complex forging part “Support” with the differences of the cross-sectional was carried out in the package Deform 3D. Options without preforming, with the implementation of a preforming upsetting of billet by elongated convex dies with eccentricity of load, and multipart closed-die forging was considered. It was found that the introduction of technological options with preforming upsetting of billet by elongated convex dies with eccentricity of load is reducing the equivalent (von Mises) stresses by 5% and the equivalent deformations by 25% in the final die impression, allowing to make lower the waste in metal flash up to 22%, and reducing the forging force more than in 4 times.

Keywords Upsetting · Preforming · Closed-die forging · Convex dies · Complex forgings · Forging force

1 Introduction

In modern mechanical engineering and metallurgical industries sufficient attention is paid to the issues related to the modernization of the old processes and the development of new technological processes, which can achieve a correlation of occupation safety issues and maximum energy and resource saving with improvement of indices of product quality [1–4]. In hot bulk closed-die forging the maximum approximation of the workpiece shape to the form of a finished product, that is preliminary forming, can significantly reduce the loss of material into flash, reduce expenditures for subsequent machining, as well as increase the resistance of forging tools [5–11]. Issues

V. Kukhar · E. Balalayeva (✉) · Y. Sahirov · A. Prysiaznyi · O. Anishchenko
Pryazovskyi State Technical University, Universytetska str., 7, Mariupol 87555, Ukraine
e-mail: balalaeva@uaeu@gmail.com

S. Hurkovska · O. Markov
Donbas State Engineering Academy, Akademichna str., 72, Kramatorsk 84313, Ukraine

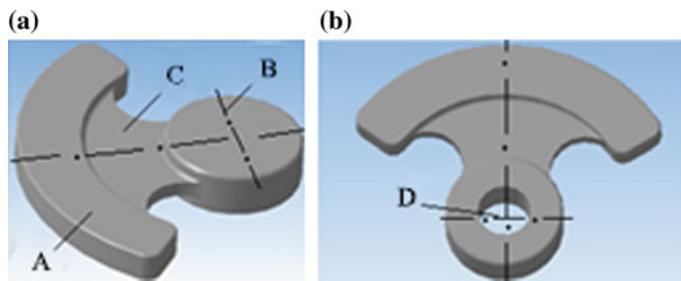


Fig. 1 Elements of the figure described in the caption should be set in italics, in parentheses, as shown in this sample caption Forging “Support” (a) and a part with a hole (b): (A) supporting portion; (B) coupling (fixing) component; (C) transient portion; (D) the hole in the part

of workpiece preforming are the most relevant for these technologies, since dies subjected to intensive impact and abrasion wear, high temperatures and pressures, run in very erroneous conditions [12–16].

In works [17–21] the authors studied and illustrated the advantage of using upsetting by flat and shaped dies as a preforming operation before the subsequent forging. However, the nomenclature of currently forged semi-finished products produced by this operation is rather limited, that requires the selection of suitable die configuration and the development of technological transitions taking into account the equipment used. The study of stress-and-strain state and forming at different variants of upsetting by shaped dies (including convex radial dies with eccentricity of load) was performed by the authors in papers [22]. One of the forgings, which forging technology improvement is expedient by the usage of workpiece preforming by upsetting by convex die with eccentricity, is a product “Support” (Fig. 1).

The objective of this work is the development and analysis of options for closed-die forging a complex forging (“Support”) with irregularities of cross section with the usage of analysis in finite element package Deform 3D and determination the technique which allows to maximum material savings at minimum forging force.

2 Material of Research

The specified forging is used in mechanical engineering designs. It is experiencing the strain of bending and stretching. Material—Steel 40 ($C = 0.37\ldots0.45\%$, $Si = 0.17\ldots0.37\%$, $Mn = 0.5\ldots0.8\%$, $Ni < 0.25\%$, $S < 0.035\%$, $P < 0.035\%$, $Cr < 0.25\%$, $Cu < 0.3\%$, $As < 0.08\%$, $Fe \sim 97\%$) GOST 1050-88, the weight of part—3.188 kg, the weight of forging (with machining allowances and overlap of hole designated in accordance with GOST 7505)—3.4 kg. Die forging of such forgings can be made on the steam hammer and crank press for hot closed-die forging. Depending on the variants of production and the equipment used, it is possible to use multiple-impression forging methods. Variants for closed-die forging on steam hammer haven’t been con-

sidered as less rational, from the point of view of material consumption as well as in terms of workshop layout, equipped with crank press for hot forging. Three variants for producing the “Support” forging were analyzed: (1) the workpiece is subjected to forging without preforming; (2) the workpiece which is shaped with eccentric upsetting by convex die is subjected to forging; (3) the workpiece with radial positioning of four forgings is subjected to multipart impression forging method.

To develop technological processes of the considered forging, the diagrams of its diameters and cross sections were built (Fig. 2). All simulation variants were performed in a finite element analysis Deform 3D package. Model of hardening, as well as the boundary conditions, was taken in accordance with the terms proposed by the program. As temperature conditions taken the isothermal deformation with temperature 1100 °C (induction heating). The models were simulated in the software package KOMPAS—3D. The first variant of the production of “Support” forging (Fig. 3) includes cutting off the rod with cylindrical billet dimensions $\varnothing 90 \times 113$ mm (weight 5.49 kg), heating, horizontal installation in the blocking impression, preliminary and final forging in the open groove with subsequent flash trimming in hot state at trimming press with tonnage of 6.3 MN. Forging with flash delivery from crank press for hot closed-die forging to trimming press is performed by a carrier.

The technological process was accompanied by unstable final size of forgings in the supporting portion border area due to under-forging to the desired size. This is explained by the fact that forgings “Support” with a thin blade and the cross-sectional differences require high specific strength (pressure) application on gravure. This is

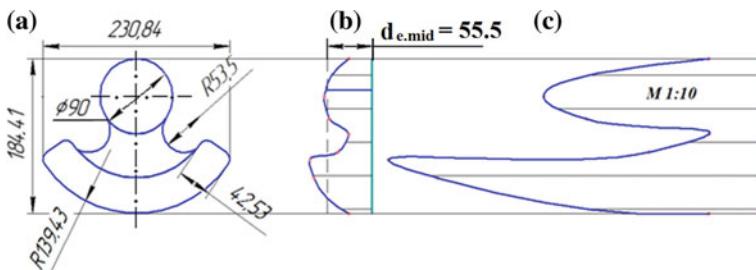


Fig. 2 Forging “Support” (a), diagram of diameters (b) and diagram of cross sections (c)

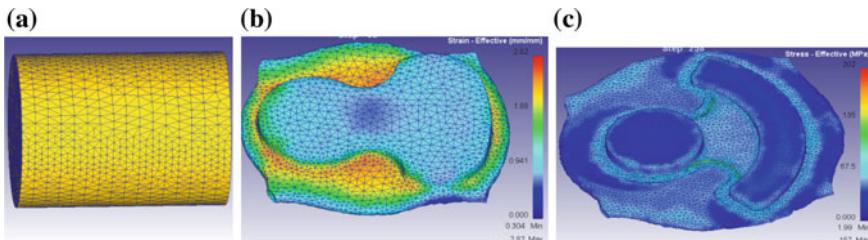


Fig. 3 Die forging process without preforming: **a** billet; **b** preforming; **c** finish-forging

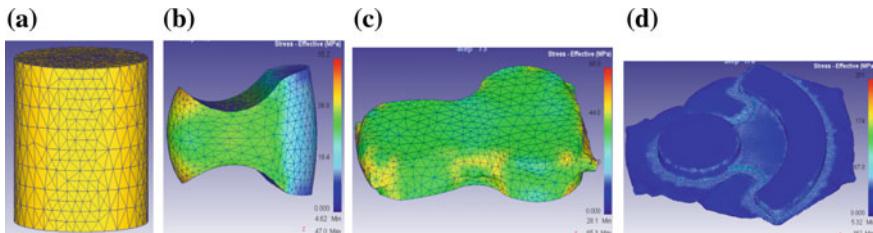


Fig. 4 Die forging process with preforming by convex dies upsetting with eccentricity of load: **a** billet; **b** shaped workpiece; **c** preforming; **d** final forging

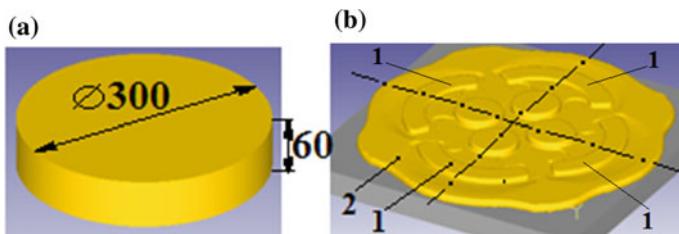


Fig. 5 Multipart forging: **a** billet for multi-forging; **b** model of multi-forging; (1) forgings; (2) flash

explained by the fact that “Support” forging with a thin blade and the cross-sectional differences require high specific strength (pressure) on the engraving to complete the final design dimensions of the product and the lack of intermediate approximation of workpiece shape to configuration of forgings leads to an unfavorable distribution of pressure on the engraving. Large reduction rate and rapid cooling of the blade of a flat forging in the die results in a significant increase in deformation resistance of the metal, which increases the proportion of elastic deformation of the press frame, and results in not finished forging and overestimated thickness of flash.

The second variant included preforming operation of upsetting by convex dies for the approximation of workpiece shape to the geometry of the blocking impression gravure (Fig. 4). This made it possible to reduce the weight of the billet from 5.49 to 4.277 kg, i.e., to achieve the metal saving of 1.21 kg, there is 22% for each forging. At the same time, partial not finished forging was excluded and full design of the blade of the forging was provided.

To localize the pressure in the area “A” (Fig. 1), to reduce the elastic deformation of the frame press, and to improve filling die impressions. It was suggested to perform the upsetting of the billet by convex elongated (radial) plates with providing of eccentricity of convex axle of plates to the vertical axle of the workpiece (Fig. 4). Similar preforming allows to redistribute the amount of metal on the estimated length of the workpiece (diagram of diameters) and to provide the distribution of metal volumes between the areas which form areas “A”, “B”, and “C” (see Fig. 1).

The third variant is multipart forging of a billet with a diameter of 300 mm and the height of 60 mm in a die at a cruciform arrangement of impressions (see Fig. 5).

Fig. 6 The graph of force required to final forging of «Support» forgings: (1) multipart forging; (2) forging without preforming; (3) forging with application of preforming by convex dies upsetting

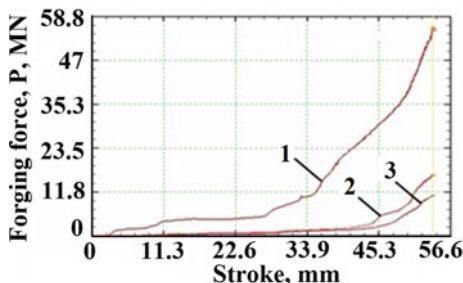


Table 1 Technical and economic indices

Indices	Variant of closed-die forging process				
Weight of a billet (Mz) (kg)	5.49	4.277	77.9	32.39	590
Number of operations (n)	3	4	133	2	66.6
Number of forgings in final impression (m), amount	1	1	100	4	400
Required force (P), MN	13.7	11.5	83.9	57	416

Finite element simulation of the plate forging in the final stage of forging with the scheme of measuring the distribution of the equivalent stress-strain indices in the longitudinal and cross section [18].

The graphs show the drop of equivalent stress indices and strain rate in the transition area of the forging (cross section (A-A)) by 8% final forging by convex plates with eccentricity of load as preforming operations. In the section (B-B) one can see the increase in the equivalent stress at 5% and equivalent strain at 25% in the place of flash formation in forging with preforming by convex plates with eccentricity of load variant, which means better filling of the cavity impression with metal. In the case of multi-forging, on the final stage of forging the highest equivalent strain is observed in the coupling section of the forging of 26 mm/mm, and the equivalent stress in the supporting portion is 74 MPa. According to the simulation results the graph of required force for forging “Support” article according to three variants was built (Fig. 6).

The graph (Fig. 6) shows that the maximum force of 57 MN for the final forging is necessary in multipart forging, and the minimal force is 11.5 MN in preforming by upsetting with convex dies with the eccentricity of load. According to the results, a table of technical and economic indices of “Support” forging was presented (Table 1).

Thus, closed-die forging with preforming has advantages both in terms of metal economy for die forging, and in the usage of less powerful and more energy efficient forging equipment.

3 The Conclusions

The analysis of the technological transition and technical and economic indices of three variants of forging a flat forgings “Support” with the presence of the cross-sectional differences (material—steel 40 GOST 1050-88) in the package Deform 3D without preforming, with the implementation of a preforming by upsetting with convex elongated dies with eccentricity of load and multipart forging was carried out. As a result of the simulation, it was revealed that the implementation of a preforming by upsetting with convex dies with eccentricity of load before the forging operation improves the filling of die gravure and reduces metal waste flash by 22%. The implementation of upsetting by convex dies with the eccentricity of load as preforming operation before the subsequent closed-die forging made it possible to reduce required forging force by 17%. Metal saving when using preforming was 22% of the base technology, the cost of forces is reduced by 16% of the base technology, and the number of operations is increased by one unit. Multipart forging modes require more force costs in contrast to forging with preforming. Therefore, the cost to equipment increases, but in the final impression it is obtained in 4 times the forged parts more than it is produced with the base technology.

References

1. Kukhar, V., Yelistratova, N., Burko, V., Nizhelska, Y., Aksionova, O.: Estimation of occupation safety risks at energetic sector of iron and steel works. *IJET (UAE)* **7**(2.23), 216–222 (2018)
2. Orlov, G.A., Orlov, A.G.: Qualimetry rating of hot-rolled pipes. *SPP* **284**, 1349–1354 (2018)
3. Chernenko, S., Klimov, E., Chernish, A., Pavlenko, O., Kukhar, V.: Simulation technique of kinematic processes in the vehicle steering linkage. *IJET (UAE)* **7**(4.3), 120–124 (2018)
4. Skubisz, P., Zak, A., Burdek, M., Lisiecki, L., Micek, P.: Design of controlled processing conditions for drop forgings made of microalloy steel grades for mining industry. *Arch. Metall. Mater.* **60**(1), 445–453 (2015)
5. Kukhar, V.V., Grushko, A.V., Vishtak, I.V.: Shape indexes for dieless forming of elongated forgings with sharpened end by tensile drawing with rupture. *SPP* **284**, 408–415 (2018)
6. Todic, V., Tepic, J., Kostelac, M., Lukic, D., Milosevic, M.: Design and economic justification of group blanks application. *Metalurgija* **51**(2), 269–272 (2012)
7. Nye, T.J.: Stamping strip layout for optimal raw material utilization. *J. Manuf. Syst.* **19**(4), 239–248 (2000)
8. Yang, Y.H., Liu, D., He, Z.Y., Luo, Z.J.: Optimization of preform shapes by RSM and FEM to improve deformation homogeneity in aerospace forgings. *CJA* **23**(2), 260–267 (2010)
9. Tang, Y.-C., Zhou, X.-H., Chen, J.: Preform tool shape optimization and redesign based on neural network response surface methodology. *Finite Elem. Anal. Des.* **44**(8), 462–471 (2008)
10. Dziubinska, A., Gontarz, A.: Limiting phenomena in new forming process for two-rib plates. *Metalurgija* **54**(3), 555–558 (2015)
11. Tofil, A., Pater, Z.: Overview of the research on roll forging processes. *ASTRJ* **11**(2), 72–86 (2017)
12. Efremenko, V.G., Chabak, Yu.G., Lekatou, A., Karantzalis, A.E., Efremenko, A.V.: High-temperature oxidation and decarburization of 14.55 wt pct Cr-cast Iron in dry air atmosphere. *Metal. Mat. Trans. A* **47A**(2), 1529–1543 (2016)

13. Hawryluk, M.: Review of selected methods of increasing the life of forging tools in hot die forging processes. *ACME* **16**, 845–866 (2016)
14. Kukhar, V., Balalayeva, E., Nesterov, O.: Calculation method and simulation of work of the ring elastic compensator for sheet-forming. In: *MATEC Web Conference*, vol. 129, p. 01041 (2017)
15. Grushko, A.V., Kukhar, V.V., Slobodyanyuk, YuO: Phenomenological model of low-carbon steels hardening during multistage drawing. *SSP* **265**, 114–123 (2017)
16. Xu, W., Li, W., Wang, Y.: Experimental and theoretical analysis of wear mechanism in hot-forging die and optimal design of die geometry. *Wear* **318**, 78–88 (2014)
17. Kukhar, V., Artiukh, V., Prysiazhnyi, A., Pustovgar, A.: Experimental research and method for calculation of ‘upsetting-with-buckling’ load at the impression-free (dieless) preforming of workpiece. In: *E3S Web Conference*, vol. 33, p. 02031 (2018)
18. Kukhar, V.V., Nikolenko, R.S., Burko, V.A.: Analysis of die-forging variants of geometrically complex forging in deform 3D package. *MMI* **1**, 18–24 (2016)
19. Meng, F., Labergere, C., Lafon, P.: Methodology of the shape optimization of forging dies. *Int. J. Mater. Form.* **3**(Suppl. 1), 927–930 (2010)
20. Puzyr, R., Kukhar, V., Maslov, A., Shchipkovskiy, Y.: The development of the method for the calculation of the shaping force in the production of vehicle wheel rims. *IJET (UAE)* **7**(4.3), 30–34 (2018)
21. Kukhar, V., Balalayeva, E., Prysiazhnyi, A., Vasylevskyi, O., Marchenko, I.: Analysis of relation between edging ratio and deformation work done in pre-forming of workpiece by bulk buckling. In: *MATEC Web Conference*, vol. 178, p. 02003 (2018)
22. Kukhar, V., Artiukh, V., Butyrin, A., Prysiazhnyi, A.: Stress-strain state and plasticity reserve depletion on the lateral surface of workpiece at various contact conditions during upsetting. *AISC* **692**, 201–211 (2018)

Optimization of DC-DC Converters for Off-Grid Lighting in Trains Using Artificial Neural Networks



Monu Malik and Ratna Dahiya

Abstract The paper manifests the optimization of DC-DC converters for off-grid lighting in trains with Artificial Neural Networks using distinct soft computing techniques as Hopfield, recurrent networks, vector quantization, cascade forward and Elman backpropagation techniques. These techniques are user-friendly and give optimum results in less time with minimum errors.

Keywords Optimization · DC-DC converters · Off-grid lighting · Trains · Artificial neural networks

1 Introduction

Due to rapid depletion of fossil fuels, there has been an increased importance of shifting to alternative sources of energy like wind energy, geothermal energy, and solar energy. This will not only help in avoiding energy crisis but will also reduce the carbon footprint and preserve ecology. In today's world, energy crisis has compelled high-end eco-architects to the modern handyman to embrace technology & engineering know-how for harnessing Solar Green Energy by usage of evolving technology and innovative implementation. Photovoltaic (PVs) are foremost familiar methods of producing electricity. Photovoltaic system makes use of solar panels each embraced with solar cells which lead to electrical energy generation. PV instatement may be on wall-mounted, rooftop or ground mounted. This PV energy creates Direct Current (DC) current from the sunlight which further finds use in many applications like charging of batteries. Solar PV system is a trump card for energy sources as

M. Malik (✉)

Department of Electronics and Electrical Engineering, Maharaja Surajmal Institute of Technology, New Delhi, India
e-mail: msmalikmonu@msit.in

R. Dahiya

Department of Electrical Engineering, National Institute of Technology, Kurukshetra, India
e-mail: ratna_dahiya@nitkkr.ac.in

once inducted it will conjure no pollution and no gas emission, it represents sailing availability to the power needs of now a day. If p and n type semiconductors are placed together one above the another and sunlight is allowed to fall on them then movement of electrons and holes takes place within the conductor and fabricates an electric current on the surface which further results in the flow of electric current.

Indian Railway has also implemented such technology usage like provision of Solar Geysers, Solar stations, Solar Street lights for Level crossings, Solar glow sign boards. In furtherance to above, the harnessing solar energy for coach electrics was experimented viz-power gear, control gear, and lighting application in the coaches. Initiative, first of its kind in the world brought results and entire Himalayan Queen (52455/56) rake plying in UNESCO heritage KLK-SML section was converted to solar power train and started on 15/06/2012. Lot of engineering and technological efforts especially in design of complete system engineering, dynamic balancing of coaches, complete system integration, reliability enhancement has gone into this project to make it a success with support of CEGE, CEE & ML. In the present study it was conceptualized to advance this step in the right direction by improving the Off-Grid Lighting of Trains. Many alternatives were explored like development of auto panel alignment in the direction of sun to enhance efficiency, etc. but it was concluded that it will be more beneficial if research is carried out in improving the DC-DC converter for the Photovoltaic array since they are not mounted on stationary structure and the train is always moving.

2 Literature Survey

Previous research work done in the field of improvement of efficiency of PV arrays was studied so as to avoid the time and effort spent in reinventing the wheel. Zhang et al. have proposed half-bridge with current fed system and full-bridge with voltage fed systems that later can employ for both fuel cell or PV systems. The paper exhibits innovative idea of switching done by zero-voltage (ZVS) which can be used in hybrid energy systems, with isolated DC-DC Converters for, e.g., in a fuel cell or solar cells and system made from superconductors [1]. Ribeiro et al. have proposed a fault-detecting technique for a photovoltaic DC-DC Converter using an fault detection with open circuit accompanied with fault-tolerant methodology for a boost converter of three level in a solar power system with batteries used as energy accumulating devices [2].

Chew et al. have proposed a 400 nW DC-DC Converter having three-output for two input of buck-boost type using MPP spooring for inside solar energy gathering. It is a single inductor-based controller. Energy gathering permits the faraway sensors of the wireless sensor network to procure the power from the environment for their lifetime [3]. Scarpa et al. have achieved higher power density in photovoltaic applications. This work explains the how combination of maximum ripple in the current can be used to obtain power density in converters. In addition to the new semiconductor devices, this technique represents less switching losses, e.g., switching losses that occur when

frequency increase [4]. Gu et al. have proposed a transformer type based on DC-DC converter working on ZVS/ZCS for solar energy microinverters [5]. Vekhande et al. have proposed a process of using DC-DC converter for the combination of photovoltaic energy with Direct Current microgrid. An Alternating Current (AC) distribution system is typical in miscellaneous commercial and industrial installations [6].

Dousoky et al. have proposed MPPT with sensorless current of DC-DC boost converter for PV battery chargers [7]. Purhonen et al. have drafted a passive component circuitry using step-up DC-DC converter in Applications [8].

Pecelj et al. have modeled a flying inductor DC-DC converter having soft switching ability acceptable for photovoltaic panels [9]. Jianwu et al. have designed a switch that can be employed on PV system. It is an isolated DC-DC converter where maximum irradiation from the sun is required or high gain from the voltage side is demanded. Previously the techniques were either switches or converter based which were used to maximize the efficiency of the system [10]. Kryukov et al. have implemented a solar energy power acclimation system with module modeling DC-DC converters for residential purposes [11].

Bansal et al. have designed a full-bridge topology for PV cells using DC-DC Converter. For the application in Power conditioning system the PV system with DC-DC Converter plays a pivotal role [12].

Bhatnagar et al. have developed many control methodologies on SIMSCAPE for the examination of PV system with Converters of DC-DC type [13].

Masri et al. have designed a DC-DC buck converter for solar energy supplication by employing Power Electronic Simulation (PSIM 9.0) [14]. Evran et al. have developed a DC-DC converter using Z source as input connected with an inductor with elevated step-up ratio acceptable for photovoltaic operation [15]. Poshtkouhi et al. have designed, sub-string MPPT module of PV so that results and performance of the system can be viewed or changed on yearly basis. In this virtual connections are done on PV cells [16]. Kim et al. have proposed the controlling strategies of resistive load connected with DC-DC converters at offset voltages in photovoltaic entreaty. Photovoltaic (PV) systems are modeled in such a way that they must conserve the steady conditions of operation at the maximum power point (MPP) despite of atmospheric conditions [17].

Das et al. have developed a framework with some degree of realism of DC-DC converter, having minimum solar voltage as input and the output have high intensity of efficiency and gains of the system [18]. Bin et al. have proposed hybrid transformer for PV module. In this boosting ratio should be high then only high efficiency from PV module can be obtained [19]. Ilango et al. have developed a DC-DC converter with hybrid PV system with high conversion ratio and is also battery powered so as to decrease the stress in the switches using MATLAB/SIMULINK [20].

Bilsalam et al. have presented a wide study analysis of PV cells with DC-DC converters such that maximum power point of photovoltaic cells can be achieved [21]. Poshtkouhi et al. have used DC-DC converter of single inductor with many inputs for calculating MPPT in photovoltaic applications when arranged in parallel connection [22]. Chudjuarjeen et al. have done network modeling of PV system using

a DC-DC boost converter [23]. Suwannatrat et al. have proposed the application of phase shifted full-bridge DC-DC converter for determining maximum power point tracking by inc conductance method [24]. Hussein et al. have shown the application of ANN in speed control of motors and those motors which are excited by some another DC source. In the study, the motor is driving a load pump which is centrifugal by operation and connected to the PV generator with buck-boost converter. Maximum Power Point is also tracked by ruling over the converter duty cycle. In this application for the speed control for DC motor with PV system the controller has many objectives. As in this research, MPP is used and then system equations approximation is done. Separately Excited DC motor characteristics were studied and transient response w.r.t voltage current and speed for step change is represented [25].

Di Piazza et al. have compared the real time simulation using ANN with analytical data interpretation of a generator excited by solar power and then controlled with DC-DC converter. The paper shows the V-I characteristics of PV field which are acquired from either numerical-based model or analytical model. The excitation to the DC-DC buck converter is provided by solar irradiance and temperature. A comparative study is also done with the previous works. (1) High quality prototypes are used in the experiments. (2) Lesser number of neurons. (3) Fast training of data. (4) Constant of PV cell should be mentioned for better results. (5) Temperature versus irradiance curve is plotted by LSR. (6) Mathematical model for discrete form can be introduced [26].

Jiteutragool et al. have done the implementation of backpropagation ANN on PV arrays for controlling the power of DC-DC boost converter. In this the input power of boost converter is varied by changing the temperature and irradiance and the required output is acquired from the boost converter. This application is best suited for battery charging. By using this methodology, fast response, less errors, speedy settling time, and minimum overshoot were procured. The PV model used in the system has the following specifications. For Simulations tool, MATLAB was operated. For this commercially convenient PV array specification: Kyocera KD135GX-LP. Experimental conditions involve optimum temperatures range between 10 and 45 °C and irradiations is maintained minimum to 200 W/m² to the maximum value up to 1000 W/m². Voltage range minimum of 12.0 V to the maximum of 24.0 V and 50 Ω (resistive load), the output power is obtained considering the step variations of 0.2 W or less than 0.2 W. The system owes speedy settling time of 0.0064 s with approximately 0.625% of voltage ripples. Varying environment conditions [27]. A glimpse from the literature view tells us about the paradigm shift from using fossil nonrenewable fuels in all areas of energy consumption and there are efforts to use these especially in the railways in recent times. To enhance the efficiency of solar panels for solar lighting other mechanical techniques may be less productive since the panels are installed on a moving train which has to face rough environmental conditions. A lot of research is underway for enhancing the efficiency of DC-DC converters for photovoltaic application of different types, i.e., buck, boost, buck-boost, etc.

Recently new soft computing techniques are being applied to enhance the efficiency and reduce the ripple of the DC-DC converters like artificial neural networks. No research has been conducted to analyse the suitability of different architectures

of artificial neural networks to enhance the accuracy of output of DC-DC converters by reducing the ripple and enhancing the efficiency. Thus, there is definitely a need to optimize DC-DC converters for off-grid lighting in trains using Artificial Neural Networks.

3 Simulation Details

The objective of the simulation work is to make a model for output of DC-DC converters for off-grid lighting in trains and choose the best suitable configuration using data mining program for extracting the system parameters of DC-DC converters for off-grid lighting in trains by employing contrasting training methodologies of artificial neural network tool box of MATLAB. A train of output of DC-DC converter has been fabricated. MATLAB has been used to simulate ANN, vigorous efforts have been laid down to enlarge the reliability of the output of the converter by operating ANN techniques on the system. The input to the training system of NN is the output of the converter which is placed in the training segment so that error-free output from the converter can be obtained. The necessary output was fed to the neural network and is fed to the individual layers of the network until the required output was received at the output side, the system goes on training the data at many iterations so that valid results should be prevalent for the output side of the converter. The obtained results then matched with the required results and then the exact picture will depict us the execution of the DC-DC converter as revealed in Fig. 1 [28].

In this arrangement, the DC-DC converter schematic is conveyed in Fig. 2. To enhance the efficiency a digital PWM has been used instead of a normal PWM as viewed in Fig. 3.

The Digital PWM has been enhanced using ANN so as to increase the accuracy of the output thereby increasing the efficiency and also reducing the ripple. Independent designing techniques of ANN are commissioned to procure the accurate output of DC-DC converter For simulation work IBM PC dual core 2 GB, 30 GB, 14" LCD,

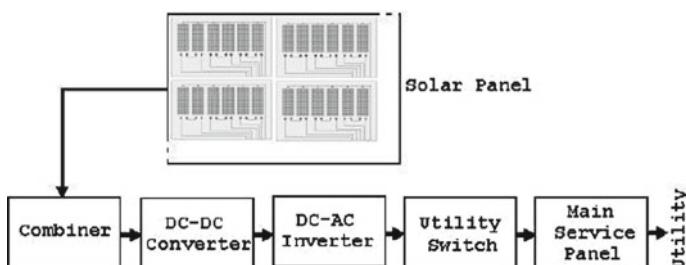


Fig. 1 Block diagram of DC-DC converter for trains

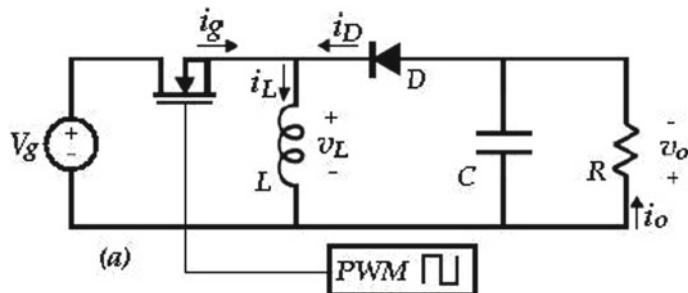
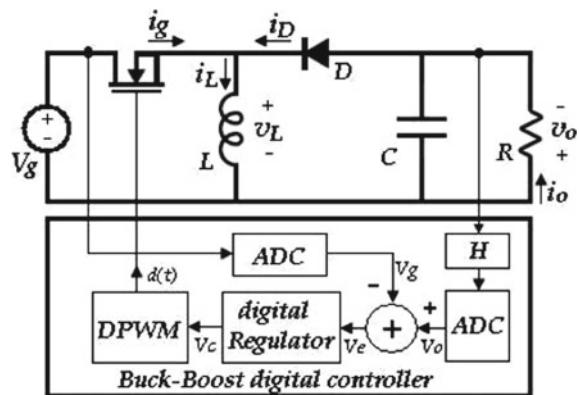


Fig. 2 Diagram of DC-DC converter

Fig. 3 Block diagram of DC-DC converter using ANN



101 keys keyboard, ps2 mouse with Matlab 7.5 having Neural Network toolbox and Simpower have been used.

4 Results and Discussions

The results are in form of traces of accuracy versus the architecture and time taken to achieve desired accuracy versus the architecture. Figure 4 shows the progress of neural network performance and training. A snapshot of the progress is reflected in framework where the block diagram of the architecture NN along with the performance with successive iterations is visible.

The error histogram of the training of neural network is seen in Fig. 5. It is evident from the figure that the accuracy is fast converging towards zero errors. The plot of Mean square error for checking the validation and study of neural network is shown in Fig. 6 and the matrix of confusion is represented in Fig. 7.

The graphical representation of gradient mu and Validation checks for different epochs is given in Fig. 8.

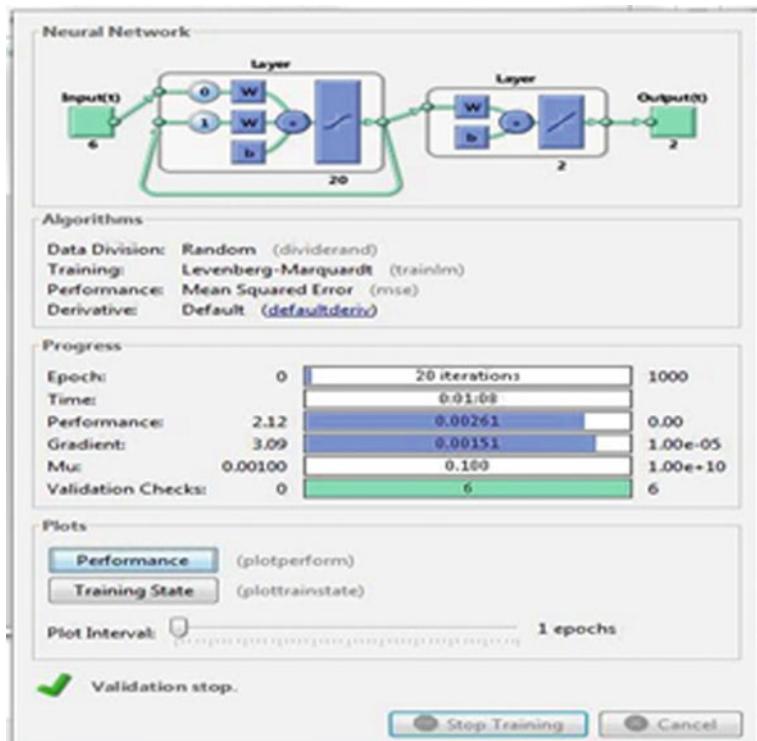


Fig. 4 Progress of neural network performance and training

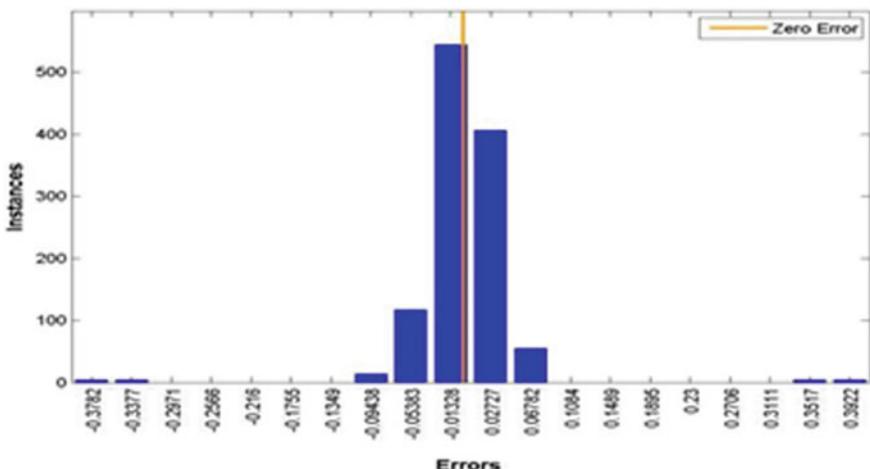


Fig. 5 Error histogram of neural network

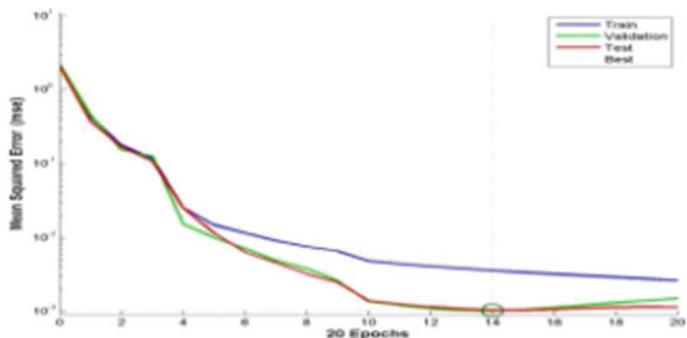
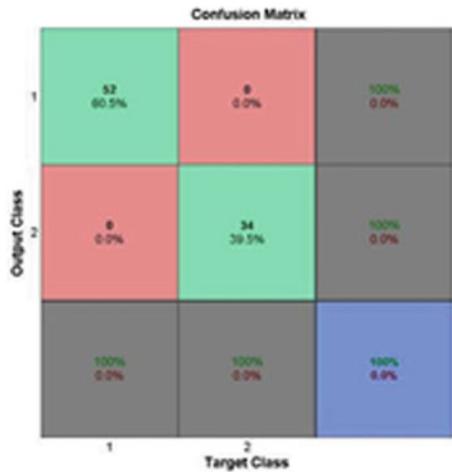


Fig. 6 Mean square error of neural network

Fig. 7 Confusion matrix of neural network



The plot in Fig. 9 shows the accuracy of diverse planning and training methods of ANN. In this study, various training on the data was done with different training techniques of Neural Network so as to establish the accuracy of the network with minimum time taken by the system from the individual data and best optimum results are obtained.

It can be seen that the time taken is least for the feed-forward input delay back-propagation network and learning vector quantization network while accuracy is high for Hopfield network and feed-forward input delay and backpropagation network and regression network as plotted in Fig. 10 while Fig. 11 represents the weight parameters of ANN structure.

It can be seen that time taken is least for feed-forward input delay backpropagation networks and learning vector quantization networks.

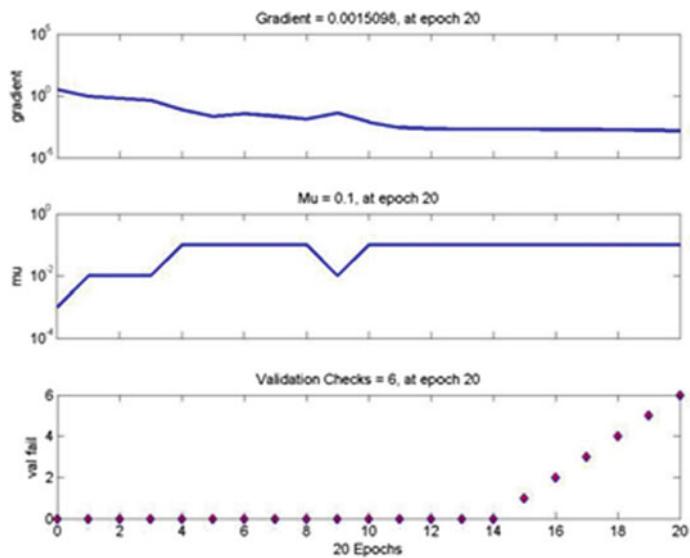


Fig. 8 Plot of gradient Mu and validation checks

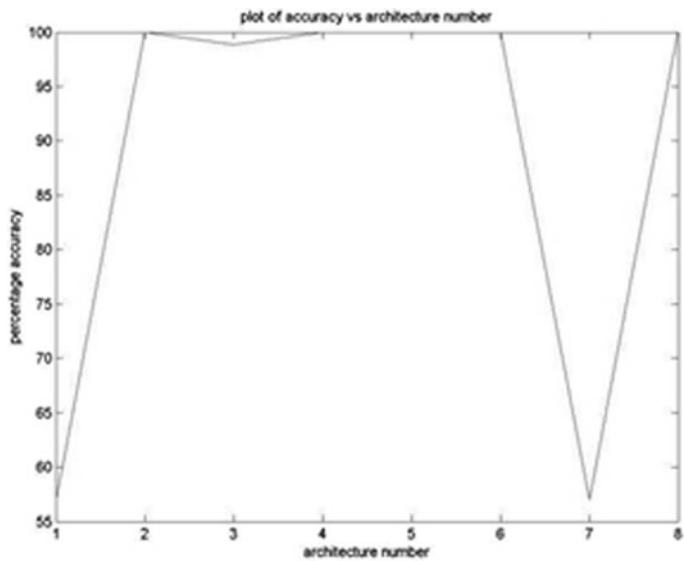


Fig. 9 Accuracy of various neural network architectures

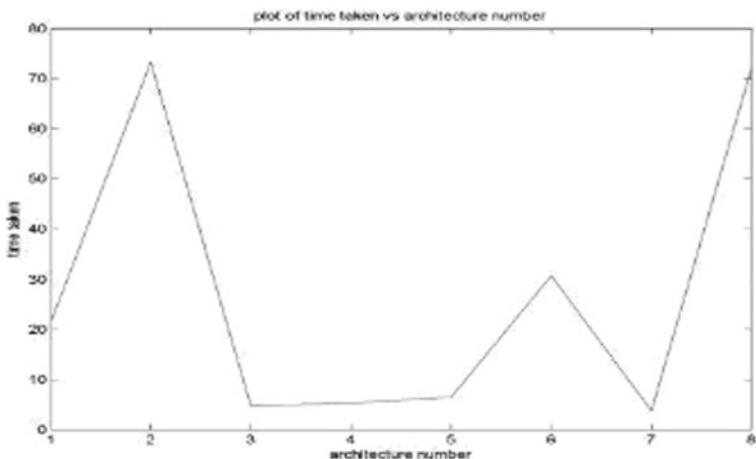


Fig. 10 Time taken to achieve accuracy by different architectures of ANN

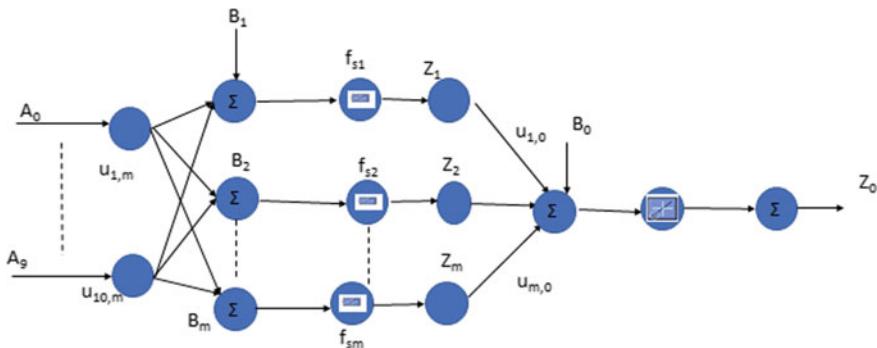


Fig. 11 Different weight parameters of ANN

5 Conclusion

It is deduced from the simulation outputs that the maximum accuracy can be achieved for DC-DC converters for off-grid lighting in trains in least possible time using for feed-forward input delay backpropagation architecture of Artificial Neural Networks.

References

1. Zhang, Z., Thomsen, O.C., Andersen, M.A.E.: Dual-input soft-switched DC-DC converter with isolated current-fed half-bridge and voltage-fed full-bridge for fuel cell or photovoltaic systems. In: Applied Power Electronics Conference and Exposition (APEC), Twenty-Eighth Annual IEEE, pp. 2042–2047, 17–21 Mar 2013 (2013)

2. Ribeiro, E., Cardoso, A.J.M., Boccaletti, C.: Fault-tolerant strategy for a photovoltaic DC-DC converter. *IEEE Trans. Power Electron.* **28**(6), 3008–3018 (2013)
3. Chew, K.W.R., Zhuochao, S., Tang, H., Siek, L.: A 400 nW single-inductor dual-input-tri-output DC-DC buck-boost converter with maximum power point tracking for indoor photovoltaic energy harvesting. In: 2013 IEEE International on Solid-State Circuits Conference Digest of Technical Papers (ISSCC), pp. 68, 69, 17–21 Feb 2013 (2013)
4. Scarpa, V.V.R., Araujo, S.V., Sahan, B., Zacharias, P.: Achieving higher power density in DC-DC converters for photovoltaic applications. In: Proceedings of the 2011-14th European Conference on Power Electronics and Applications (EPE 2011), pp. 1, 10, 30 Aug 2011–1 Sept 2011 (2011)
5. Gu, B., Dominic, J., Lai, J.-S., Ma, H.: Hybrid transformer ZVS/ZCS DC-DC converter for photovoltaic microinverters. In: 2013 Twenty-Eighth Annual IEEE Applied Power Electronics Conference and Exposition (APEC), pp. 16, 22, 17–21 Mar 2013 (2013)
6. Vekhande, V., Fernandes, B.G.: Module integrated DC-DC converter for integration of photovoltaic source with DC micro-grid. In: IECON 2012—38th Annual Conference on IEEE Industrial Electronics Society, pp. 5657, 5662, 25–28 Oct 2012
7. Dousoky, G.M., Ahmed, E.M., Shoyama, M.: Current-sensorless MPPT with DC-DC boost converter for photovoltaic battery chargers. In: 2012 IEEE Energy Conversion Congress and Exposition (ECCE), pp. 1607, 1614, 15–20 Sept 2011
8. Purhonen, M., Hannonen, J., Strom, J., Silventoinen, P.: Step-up DC-DC converter passive component dimensioning in photovoltaic applications. In: 2012 IEEE 27th Convention of Electrical & Electronics Engineers in Israel (IEEEI), pp. 1, 5, 14–17 Nov 2012 (2012)
9. Pecelj, I., De Haan, S.W.H., Ferreira, J.A.: A soft-switched, flying inductor DC-DC converter suitable for photovoltaic panels. In: 2012 7th International Power Electronics and Motion Control Conference (IPEMC), vol. 2, pp. 1241, 1246, 2–5 June 2012 (2012)
10. Zeng, J., Qiao, W., Qu, L.: A single-switch isolated DC-DC converter for photovoltaic systems. In: 2012 IEEE Energy Conversion Congress and Exposition (ECCE), pp. 3446, 3452, 15–20 Sept 2012
11. Kryukov, K.V., Valiev, M.M.: Residential photovoltaic power conditioning system with module integrated DC-DC converters. In: 2012 15th International Power Electronics and Motion Control Conference (EPE/PEMC), pp. DS3b.9-1, DS3b.9-4, 4–6 Sept 2012 (2012)
12. Bansal, S., Saini, L.M., Joshi, D.: Design of a DC-DC converter for photovoltaic solar system. In: 2012 IEEE 5th India International Conference on Power Electronics (IICPE), pp. 1, 5, 6–8 Dec 2012 (2012)
13. Bhatnagar, P., Nema, R.K.: Control techniques analysis of DC-DC converter for photovoltaic application using SIMSCAPE. In: 2012 IEEE 5th India International Conference on Power Electronics (IICPE), pp. 1, 6, 6–8 Dec 2012 (2012)
14. Masri, S., Mohamad, N., Hariri, M.H.M.: Design and development of DC-DC buck converter for photovoltaic application. In: 2012 International Conference on Power Engineering and Renewable Energy (ICPERE), pp. 1, 5, 3–5 July 2012
15. Evran, F., Aydemir, M.T.: A coupled-inductor Z-source based Dc-Dc converter with high step-up ratio suitable for photovoltaic applications. In: 2012 3rd IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG), pp. 647, 652, 25–28 June 2012
16. Poshtkouhi, S., Biswas, A., Trescases, O.: DC-DC converter for high granularity, sub-string MPPT in photovoltaic applications using a virtual-parallel connection. In: 2012 Twenty-Seventh Annual IEEE Applied Power Electronics Conference and Exposition (APEC), pp. 86, 92, 5–9 Feb 2012 (2012)
17. Kim, K.A., Li, R.M., Krein, P.T.: Voltage-offset resistive control for DC-DC converters in photovoltaic applications. In: 2012 Twenty-Seventh Annual IEEE Applied Power Electronics Conference and Exposition (APEC), pp. 2045, 2052, 5–9 Feb 2012 (2012)
18. Das, M., Agarwal, V.: A novel, high efficiency, high gain, front end DC-DC converter for low input voltage solar photovoltaic applications. In: IECON 2012—38th Annual Conference on IEEE Industrial Electronics Society, pp. 5744, 5749, 25–28 Oct 2012

19. Gu, B., Dominic, J., Lai, J.-S., Zhao, Z., Liu, C.: High boost ratio hybrid transformer DC-DC converter for photovoltaic module applications. In: 2012 Twenty-Seventh Annual IEEE Applied Power Electronics Conference and Exposition (APEC), pp. 598,606, 5–9 Feb 2012
20. Ilango, K., Ram, S., Nair, M.G.: A hybrid photovoltaic-battery powered DC-DC converter with high conversion ratio and reduced switch stress. In: 2012 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES), pp. 1, 5, 16–19 Dec 2012 (2012)
21. Bilsalam, A., Boonyaroonate, I., Haema, J., Chunkag, V.: A study effect and improved maximum power point of photovoltaic cells with DC-DC converter. In: 2012 Asia-Pacific Power and Energy Engineering Conference (APPEEC), pp. 1, 4, 27–29 Mar 2012 (2012)
22. Poshtkouhi, S., Trescases, O.: Multi-input single-inductor dc-dc converter for MPPT in parallel-connected photovoltaic applications. In: 2011 Twenty-Sixth Annual IEEE Applied Power Electronics Conference and Exposition (APEC), pp. 41, 47, 6–11 Mar 2011 (2011)
23. Chudjuarjeen, S., Jimenez, J.C., Jayasuriya, S., Nwankpa, C.O., Miu, K., Sangswang, A.: DC-DC boost converter with network model for photovoltaic system. In: 2011 IEEE Energy Conversion Congress and Exposition (ECCE), pp. 1273, 1278, 17–22 Sept 2011 (2011)
24. Suwannatrat, P., Liutanakul, P., Wipasuramonton, P.: Maximum power point tracking by incremental conductance method for photovoltaic systems with phase shifted full-bridge dc-dc converter. In: 2011 8th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), pp. 637, 640, 17–19 May 2011 (2011)
25. Hussein, A., Hirasawa, K., Hu, J., Murata, J.: The dynamic performance of photovoltaic supplied dc motor fed from DC-DC converter and controlled by neural networks. In: Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN '02, vol. 1, pp. 607, 612 (2002)
26. Di Piazza, M.C., Pucci, M., Ragusa, A., Vitale, G.: Analytical versus neural real-time simulation of a photovoltaic generator based on a DC-DC converter. In: Energy Conversion Congress and Exposition. ECCE 2009. IEEE, pp. 3350, 3356, 20–24 Sept 2009 (2009)
27. Jiteurtragool, N., Wannaboon, C., San-Um, W.: A power control system in DC-DC boost converter integrated with photovoltaic arrays using optimized back propagation artificial neural network. In: 2013 5th International Conference on Knowledge and Smart Technology (KST), pp. 107, 112, 31 Jan 31–1 Feb 2013 (2013)
28. Sulaiman, D.R., Amin, H.F., Said, I.K.: Design of high efficiency DC-DC converter for photovoltaic solar home applications. *J. Energy* **4**(11) (2010)

A Study on Noninvasive Body Wearable Sensors



Shanu Bhardwaj and S. N. Panda

Abstract The paper furnishes a review on some of the notable research works done on the noninvasive body wearable sensors. The wearable device is embedded with sensors attached to the patient. With the help of sensors, the parameters of the patient can be transferred remotely for further analyses. Traditional devices used for monitoring consist of wires which lead to soreness and even more pain when these devices are removed out of the body. So to overcome these limitations, there exists an evolution of noninvasive body wearable sensors. This paper aims to provide that how these wearable sensors are different from the traditional devices used, show the architectural view of the wearable sensors network with the perspective of healthcare, and also discuss the performance, i.e., the comparative study of different sensors from various research papers of past years.

Keywords Wireless technology · Noninvasive · Body wearable · Embedded sensor · Physiological parameters

1 Introduction

Wearable sensors, nowadays, gain popularity in many of the real-time applications such as health care, security, commercial fields, entertainment, and many more [1]. These wearable sensors had received a great fascination over the ancient decades for monitoring the health of wearers [2]. Monitoring of parameters of the patient is the essential application of the sensors [1]. Wearable devices integrated with sensors are different from traditional devices used for monitoring purpose, for the reason that the presence of many wires leads to soreness and even more pain when removed. So there was an exigency or demand of the noninvasive wearable devices embedded with sensors [3]. In the field of health care, monitoring of physiological activities or

S. Bhardwaj (✉) · S. N. Panda
Chitkara University Institute of Engineering and Technology, Chitkara University,
Chandigarh, India
e-mail: shanu.bhardwaj@chitkara.edu.in

parameters such as temperature, ECG, brain activity, heart rate, muscle motion, and other vital parameters is possible only because of embedded sensors in the wearable devices. These advancements in the sensor and the wireless technology enable the formation of the new healthcare system for monitoring with wearable sensors.

In the modern era, still, many of the countries are suffering from a lack of skillful medical care or health staffs. On the other side, high-tech products are developed day by day. Because of these products, our day-to-day life actually becomes more efficient and convenient [4]. Moreover, wearable devices and sensors provide an innovative way to record the real-time health information from the human body securely [5].

This paper aims to provide the study of noninvasive body wearable sensors, study how these wearable sensors are different from the traditional devices used, show the architectural view of the wearable sensors network with the perspective of health care, and also discuss the performance, i.e., the comparative study of different sensors from various research papers of past years.

This paper has been organized further in the following way. Section 2 discusses the literature review. Section 3 presents the architectural view of the network of wearable sensors. Section 4 discusses the comparative study done by determining different research papers. Section 5 discusses the conclusion of the paper.

2 Literature Review

Leu et al. [4] proposed a system for monitoring the physiological data in the smartphone with the help of body sensors inserted in the “smart shirt” of an individual. Physiological parameters were required for the treatment, monitoring, and prevention of the diseases. This system collects the physiological parameters with the sensors inserted in the smart shirt. These vital signs or parameters were gathered and forwarded to the smartphone in real time. Now, with the Wi-Fi, the parameters were delivered to the cloud of remote health care and also the data of an individual were stored in the cloud, so that the authorized health staff can view the data whenever necessary [2] (Fig. 1).

Tseng et al. [6] proposed a technique to integrate signals from two sensors, SpO₂ and ECG. This proposed method has an advantage that the functioning of two sensors was included in one wearable device. This wearable device was small in size as well as cost-effective. It was allowed that the data could be stored in real time, so that the collected data can be monitored by health staff or physicians for the precision of medicine or in any emergency case [6].

Kim et al. [7] presented the advancement toward the blooming of noninvasive sensing system of glucose using wearable sensors. The existing glucose sensing system was the blood test using finger stick which was painful. The recent report declared that the glucose sensing in human improves the route toward the maintenance of diabetes as well as glycemic control. The author proposed a sticker like a sensor to be put on the arm and it detects the sweat and interstitial fluid in the

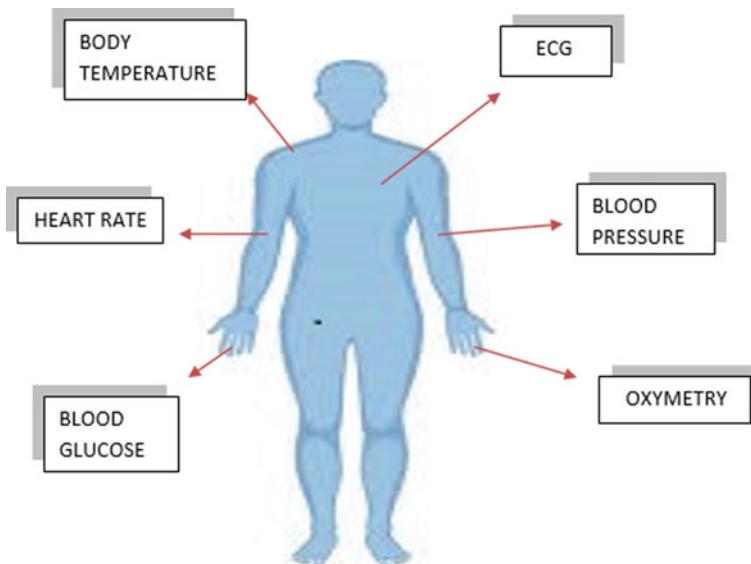


Fig. 1 Physiological parameters [2]

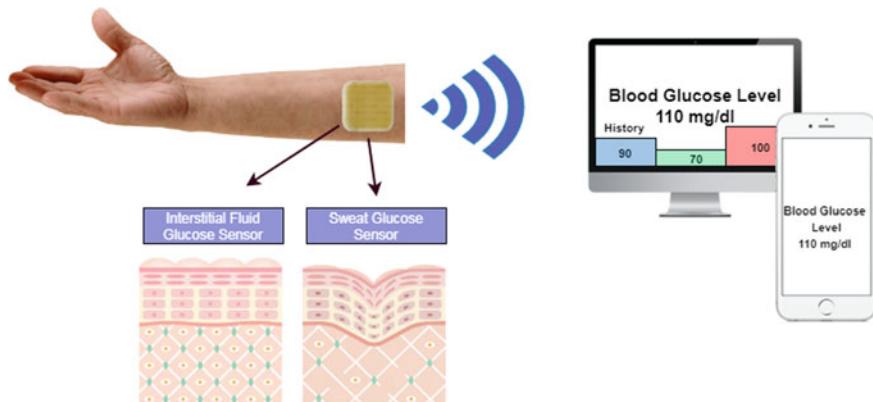


Fig. 2 Noninvasive glucose sensor [3]

glucose. With the help of the Wi-Fi module, the parameters were displayed in the smartphones or laptops [3] (Fig. 2).

Lin et al. [8] developed a system for the monitoring of heart rate via sensors of heart rate and also the Bluetooth module for data transmission wirelessly. The parameters or the signals were collected by the sensor, processed by the signal process circuit, and then transmitted with the help of Bluetooth, and the whole data was generated in the smartphone for monitoring the heart rate in real time. The author concluded that the proposed system was cost-effective for daily monitoring of heart rate [8].

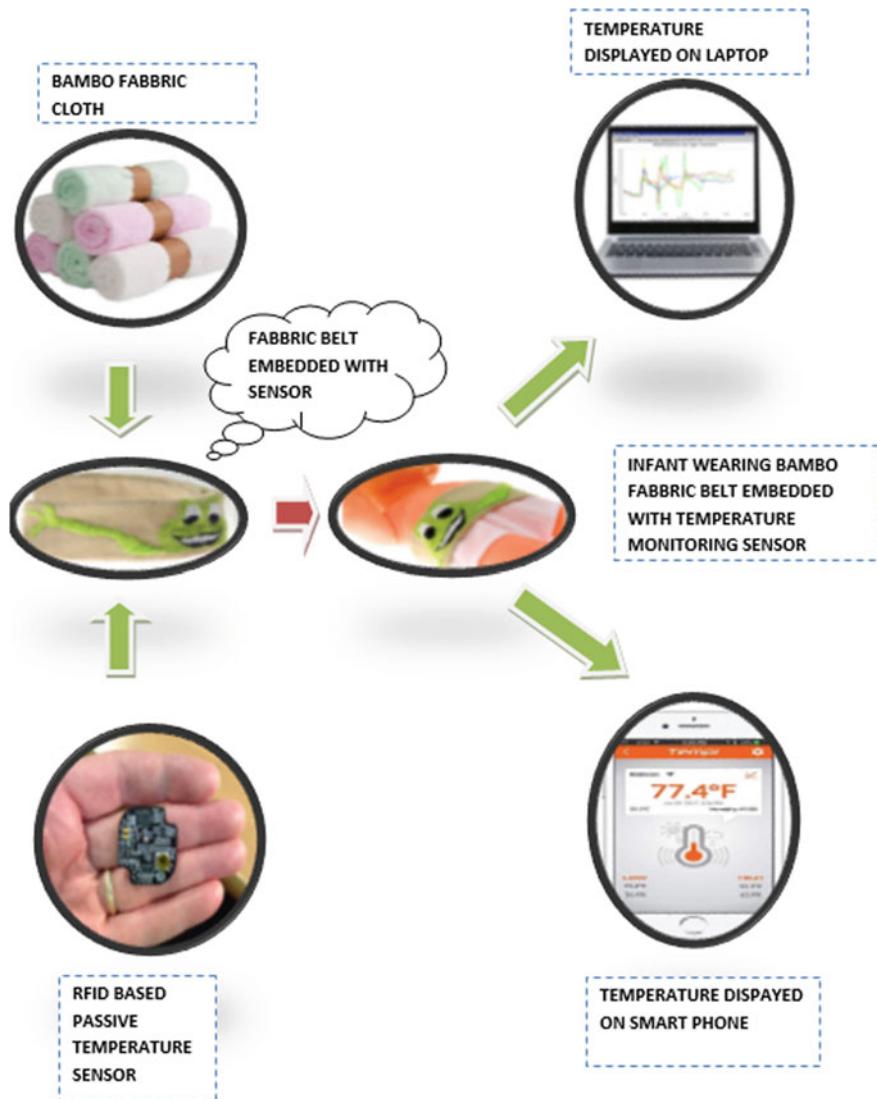


Fig. 3 Infant temperature monitoring via fabric belt embedded with sensor [4]

Yao et al. [9] discussed the opportunities, challenges, and the future scope of the nanomaterial-enabled wearable sensor which could comfortably be attached to the skin for the monitoring purpose of patient's physiological parameters over the Internet. This paper had covered the recent advancements in the nanomaterial-enabled wearable sensors including electrophysiological, temperature, and electrochemical sensors [9].

Chen et al. [10] proposed a design for infant temperature monitoring via wearable sensors. Vital parameters of infants like body temperature, the oxygen level in the blood, and ECG need to be observed for urgent medical treatment. Currently, the monitoring of body temperature is done by the sticky thermistors which consists of wires that lead to tenderness and even more pain while removing those sticky sensors. Due to this reason, the author proposed a noninvasive temperature monitoring wearable devices such as “smart belt” for the neonates or infants to provide appropriate care and treatment [3] (Fig. 3).

3 Architecture of Wearable Sensor Network

The architecture of wearable sensor network in health care is shown in Fig. 4. Monitoring the health status of the patient via a wireless sensor network has been in demand. There are various projects in health care that are in the development or implementation phase.

In Fig. 4, the sensor data can be generated from anywhere; the sensors are adhered such as remote practitioners, local practitioners, hospitals, outdoor, homecare, ambulance, and many more. The data by the sensors is routed through a local wireless gateway or remote wireless gateway and reaches to the main server. Then, the server distributes the data to the database for the retrieval of the data. Further, the data of the patient or any individual reaches to the destination, i.e., medical health care or any other, over the Internet to their smartphones and laptops.

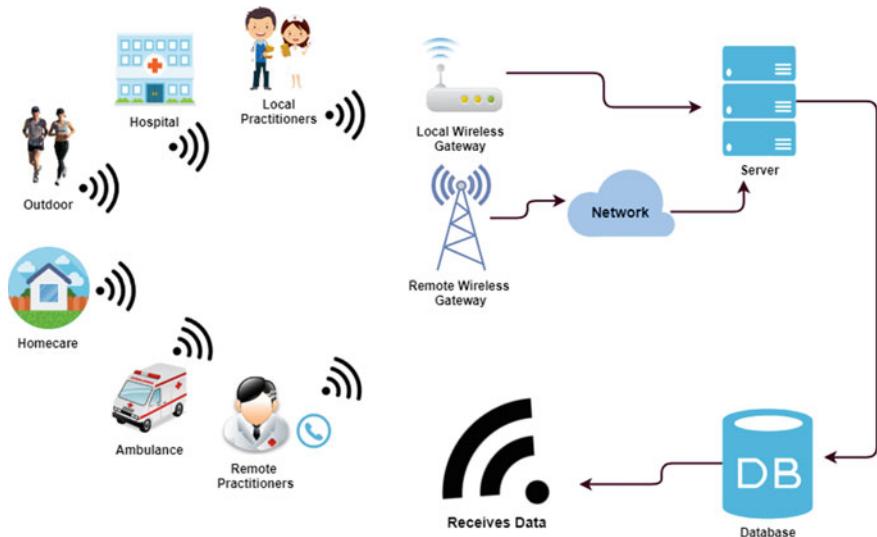


Fig. 4 Architecture of wearable sensor network

Table 1 Comparative study of data on different ECG sensors

Parameters	Leu [5]	Ameen [7]	Kim [8]	Lee [9]
Channel	1	1	1	3
Supply (V)	3	3	3.3	3
Power (mW)	12.5	84.83	115	375
Storage	SD card	SD card	N/A	SD card
Size	$5.8 \times 5.0 \times 10 \text{ cm}^3$	$5.8 \times 5.0 \times 0.4 \text{ cm}^3$	$5.5 \times 3.4 \times 1.6 \text{ cm}^3$	N/A
Weight	20	38	20.7	N/A
Electrode	N/A	FPDE	N/A	Commercial electrode

Table 2 Comparative study of two different HR sensors

Parameters	Gambi [6]	Lin [11]
Processing	Based on the processing of RGB images captured	Signals are first processed by signal process circuit
Observation	Changes are observed in the form of change in color due to blood flow	Changes are observed in the form of change in signals
Network	No wireless transmissions exist	Wireless transmissions of the signals via Bluetooth
Access	No remote access	It can be remotely accessed in real time
Size	Microsoft Kinect is expensive	Cost-effective solution

4 Performance

This section provides the performance based on the comparative study of different sensors by evaluating various parameters. Table 1 shows the comparative study of different ECG sensors on the basis of parameters such as channel, supply, power, storage, size, weight, and electrode. Similarly, Table 2 shows the comparative study of two different heart rate (HR) sensors on the parameters that are processing, observation, network, access, and size.

5 Conclusion and Future Scope

Wearable devices integrated with sensors are different from traditional devices used because the presence of many wires leads to soreness and even more pain when removed. The extent of the research work on this particular topic is increasing day by day with the growth in the market so, in concern of that, a brief review on some of the

eminent research work done on noninvasive wearable body sensors had been depicted in this paper. This paper also provides the case study of various body wearable sensors that are different from the traditional one and also does the comparative study of the various sensors in the same field, and also shows the architectural view of the wearable sensors network with the perspective of health care.

Future work comprises that these wearable sensors can be embedded within the wheelchair, so that the critical patients can get treatment well with ease. The wheelchair consists of different sensors so that the physiological parameters of the patient sitting on the wheelchair can be monitored wirelessly without soreness and pain.

References

1. Mukhopadhyay, S.C.: Wearable sensors for human activity monitoring: a review. *IEEE Sens. J.* **15**(3), 1321–1330 (2015)
2. Bandodkar, A.J., Jeerapan, I., Wang, J.: Wearable chemical sensors: present challenges and future prospects. *ACS Sens.* **1**(5), 464–482 (2016)
3. Chen, W., Dols, S., Oetomo, S.B., Feijis, L.: Monitoring body temperature of newborn infants at neonatal intensive care units using wearable sensors. In: Proceedings of the Fifth International Conference on Body Area Networks (MMC), pp. 188–194 (2010)
4. Leu, F., Ko, C., You, I., Choo, K.K.R., Ho, C.L.: A smartphone-based wearable sensors for monitoring real-time physiological data. *Comput. Electr. Eng.* **65**, 376–392 (2018)
5. Al, Ameen M., Liu, J., Kwak, K.: Security and privacy issues in wireless sensor networks for healthcare applications. *J. Med. Syst.* **36**(1), 93–101 (2012)
6. Tseng, C., et al.: Application of a minimized wearable device combined with SpO₂ and ECG sensors to detect stenosis or occlusion of arteriovenous fistula/graft, progression of arteriosclerosis and arrhythmia. In: 7th International Symposium on Next Generation Electronics (ISNE), pp. 1–4 (2018)
7. Kim, J., Campbell, A.S., Wang, J.: Wearable non-invasive epidermal glucose sensors: a review. *Talanta* **177**, 163–170 (2018)
8. Lin, Z., et al.: Triboelectric nanogenerator enabled body sensor network for self-powered human heart-rate monitoring. *ACS Nano* **11**(9), 8830–8837 (2017)
9. Yao, S., Swetha, P., Zhu, Y.: Nanomaterial-enabled wearable sensors for healthcare. *Adv. Healthc. Mater.* **7**(1), 1–27 (2018)
10. Chen, C.Y., et al.: A low-power bio-potential acquisition system with flexible PDMS dry electrodes for portable ubiquitous healthcare applications. *Sensors (Switzerland)* **13**(3), 3077–3091 (2013)
11. Miao, F., Cheng, Y., He, Y., He, Q., Li, Y.: A wearable context-aware ECG monitoring system integrated with built-in kinematic sensors of the smartphone. *Sensors (Switzerland)* **15**(5), 11465–11484 (2015)
12. Lee, H.C., et al.: An ECG front-end subsystem for portable physiological monitoring applications. In: Proceedings of the 2011 International Conference on Electric Information and Control Engineering ICEICE 2011, pp. 6359–6362 (2011)
13. Gambi, E., et al.: Heart rate detection using microsoft kinect: validation and comparison to wearable devices. *Sensors* **17**(8), 1776 (2017)

An Outlier Accuracy Improvement in Shilling Attacks Using KSOM



Anjani Kumar Verma and Veer Sain Dixit

Abstract Due to the rapid technological changes, these days collaborative filtering-based recommender systems are being widely used worldwide. Collaborative filtering approach is more vulnerable from being attacked because of its open nature. The attackers may rate the fake ratings to disturb the systems. In this paper, unsupervised Kohonen Self-Organizing Map (KSOM) clustering technique is used to make a better detection between genuine and fake profiles to reduce profile injection attacks and compared with existing techniques Enhanced Clustering Large Applications Based on Randomized Search (ECLARANS) and Partition Around Medoids (PAM) with variants of attack size. It has been noticed that KSOM outperforms over ECLARANS and PAM techniques with good outlier accuracy.

Keywords Recommender systems · Collaborative filtering · Shilling attacks · KSOM · ECLARANS · PAM · Outlier accuracy

1 Introduction

E-commerce has become the most popular way of trading on the Internet. User always wants safe trading but due to the presence of malicious users, they always feel insecure about their data. Shilling attacks are one of the major challenges these days and major concerns of researchers as well as stakeholders.

A. K. Verma (✉)

Department of Computer Science, University of Delhi, Delhi, India

e-mail: anjaniverma29@gmail.com

V. S. Dixit

Department of Computer Science, ARSD College, University of Delhi, Delhi, India

e-mail: veersaindixit@rediffmail.com

Due to the exponential increase of information on daily basis, information overloading is a major issue in recommender systems [1] used on the web [2], which is required to overcome. The databases have been filled with many anonymous ratings such as fake profiles that are hard to identify, and there is a need to use some robust techniques [3].

One of the problems of shilling attack [4] is to distort the real existence of data by any mean. In this process, there could be any set of profiles of a particular interest which would solve a purpose and if, meanwhile, someone is going to misuse the data then it will totally change the sense of it and will become very difficult to identify.

There is a need to solve the above-said problem for the sake of reality and check all those profiles that are available on E-commerce websites nowadays. It is observed that there are many profiles that could change the pattern of the data consequent to which it would be difficult to understand the information quality, hence, hampering the future prospects regarding the overall popularity of the e-commerce sites.

It is very difficult to measure humane perception in terms of their views or liking or disliking, which is the main motivation of this paper. Sometimes, it is very hard to predict the ratings of similar user by their perception, as from their valuing of recommendations with varying duration of time. Online users are frequently doing and updating the ratings but this may not sure that viewpoint expressed by the users is actual or something else [5].

The base paper [6] has described the outlier detection methods in which different approaches have been taken in the view of collaborative filtering. After taking an assumption of outlier is in small percent of the total data, the authors has taken small number of clusters that contain outliers [7]. PAM and ECLARANS both are unsupervised clustering algorithms [8] used for outlier detection [9].

The subsequent sections are written as follows:

In Sect. 2, description of related works has been detailed. In Sect. 3, the methodology has been explained. In Sect. 4, the experiments have been carried out. At the end, paper is concluded in Sect. 5.

2 Background

In this section, we have tried to come across different works done so far in this area. In [10], statistical detection techniques are described by which different attack models are proposed. Ninety percentage of precision is seen in the detection of various attack models based on PCA clustering technique.

In [11], the author proposed the method, a binary decision tree (BDT) [12], which is constructed by recursively clustering the training data to locate the fake attack profiles via bisecting k-means clustering algorithm [13]. Here, author hypothesized that internal nodes holding attack profiles demonstrate high intra-cluster correlation due to their high similarity among themselves.

In [14], attack properties on different attack models are identified using PCA-based detector and produced analytically good results. After that, they address the obfuscated attacks using multivariate Gaussian mixture model. Finally, supervised and unsupervised Neyman–Pearson detectors are designed. It outperforms over PCA-based detector.

In [15], Angle-Based Outlier Detection (ABOD) is discussed, where the angle differences have been calculated between the one vector point and other points detected as an outlier. In [16], PAM algorithm is used to obtain high accuracy in comparison of k-means.

In [6], performance-wise supremacy of ECLARANS in terms of accuracy over PAM algorithm is better. Hence, we have taken it as a problem that can look whether KSOM clustering algorithm gives the better results in accuracy over ECLARANS as well as PAM or not. We are aimed here to improve the accuracy of detection of fake profiles. Taking into consideration, we have used KSOM clustering technique and then we have compared the results obtained by processing the algorithm with the ECLARANS and PAM, respectively, on the basis of similar kind of parameters.

3 Methodology

Due to easy interpretation of data along with the reduction of dimensionality, KSOM is used. In KSOM technique, grid clustering makes it easy to observe similarities in the data. By applying some optimization techniques, KSOM can be trained in a short amount of time. KSOM [17, 18] is capable of handling several types of clustering problems while providing a useful and interactive summary of the data. After studying the KSOM, one very important limitation was found that it requires necessary and sufficient data in order to develop meaningful clusters in which weight vectors must be based on data that can successfully group and distinguish inputs. KSOM is a robust learning technique being the reason that visual inspection has to be done cautiously because the topology preservation property of the trained map has to be checked before starting visual inspection. Otherwise, visual inspection can mislead one's thought process.

3.1 KSOM Algorithm

A KSOM learns to classify the training data without any external supervision. The steps for outlier detection are as follows:

```

begin
    initialize weight vector
    t<-0
    for epoch<-1 to Nepochs do
        Interpolate new values for  $\alpha(t)$  and  $\beta(t)$ 
        for record<-1 to Nrecords do
            t<-t+1
            for k<-1 to K do
                Compute distances  $d_k$  using Eq.1
                 $d_k(t) = ||x(t) - w_k(t)||^2$ 
            end
            Compute winning node c using Eq.2
             $d_c(t) = \min(d_k(t))$ 
        end
        for k<-1 to K do
            Update weight vectors  $w_k$  using Eq.3
             $w_{ij}(\text{new}) = w_{ij}(\text{old}) + \alpha[x_i - w_{ij}(\text{old})]$ 
        end
    end
end

```

(1)

(2)

(3)

3.2 Model

While using movie recommender system based on collaborative filtering approach, the recommendations generated may be genuine recommendations or could be of malicious nature. The genuine recommendations are termed as genuine profiles, and malicious recommendations are termed as attack profiles in this paper. Both the profiles are stored in the database. Once profiles stored into the database are provided as input vectors to the KSOM, after completing the adaptation process by KSOM algorithm, clusters of different sizes are produced. With the help of outlier accuracy, the suspected profiles are obtained.

The framework of method used is illustrated in Fig. 1.

4 Experimental Results

4.1 Dataset

In this section, we have made a comparative analysis between the three clustering algorithms on the same dataset by taking the common parameter. There are three phases of implementation in which different percentages of filler items have been

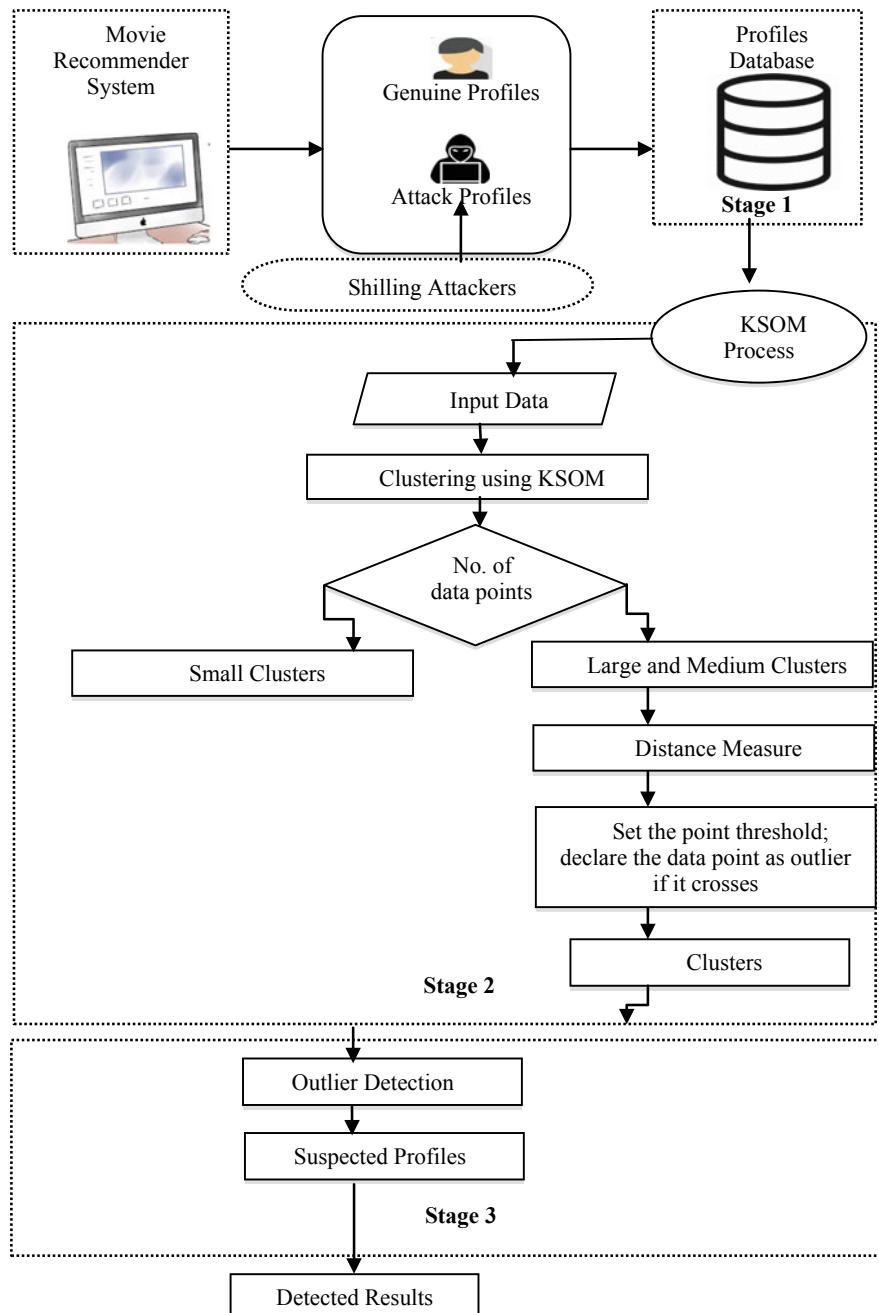


Fig. 1 Framework of model

Fig. 2 Cluster size using ECLARANS

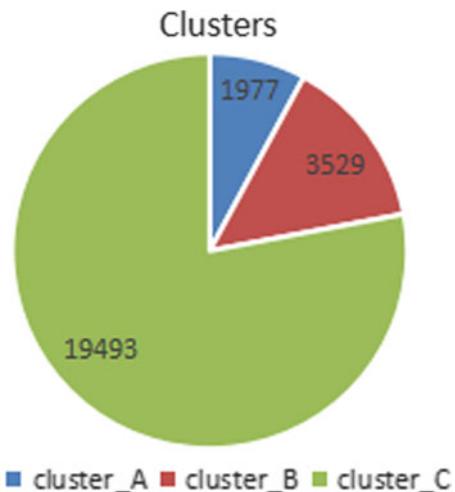


Table 1 Dataset architecture

UserId	MovieId	Rating	Timestamp
196	242	3	881250949
186	302	3	891717742
...

taken, that is, 40%, 60%, and 70%, and we have labeled them as “case1”, “case 2”, and “case 3”, respectively, in the section below. Here, we have taken three clusters for each case to analyze the performance.

Dataset input: MovieLens¹ 1 M

Actual output: Cluster (suspected profile)

Performance Evaluation Factor: Outlier Accuracy.

In this experiment, we have taken MovieLens-1 M dataset. This dataset contains 1,000,209 ratings from 6,040 users and 3,900 movies (items). The ratings are ranging from 1 to 5. The experiment has been performed on Weka [19] and MATLAB 9.0 [20] with configuration of 8 GB RAM. The architecture is shown in Table 1.

¹“MovieLens dataset” available at <https://movielens.org/> accessed on April 12, 2018.

In ECLARANS algorithm, the following size of clusters has been generated which is shown in Fig. 2. Here, we see that cluster A and cluster B are small-size clusters in Table 2. This can be seen that user is included in clusters A and B considering the attack profiles.

4.2 Experimental Setup

4.2.1 Case 1

In this dataset, filler items are 10,000 and target items are 15,000 as we have randomly taken 25,000 anonymous ratings of 547 users and 1407 movies.

In PAM algorithm, the following size of clusters was generated which is shown in Fig. 3. We identify outliers as the data objects that belong to a cluster having size lesser than half of the average number of points in the k clusters. Here, cluster A and cluster B are considered as small clusters that can be seen in Table 2.

Fig. 3 Cluster size using PAM

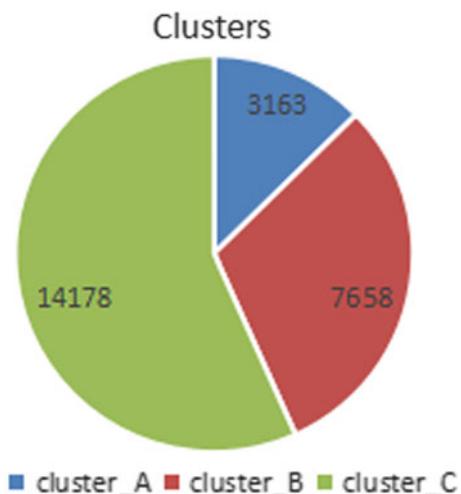
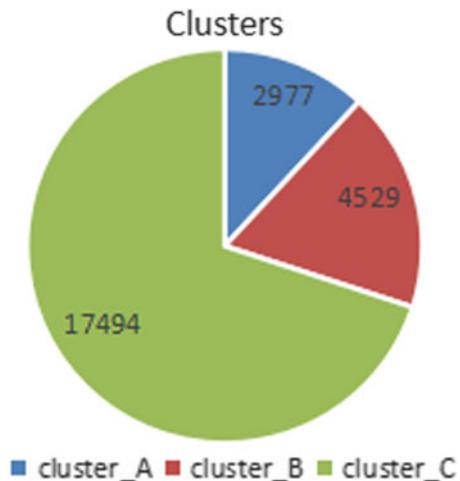


Table 2 Clusters generated from KSOM, ECLARANS, and PAM on different filler items

Algorithm	No. of clusters	40% filler items	60% filler items	70% filler items
PAM	cluster A	3163 < 8334 items	15,432 > 12,500 items	35,145 > 20,000 items
	cluster B	7658 < 8334 items	12,273 < 12,500 items	15,672 < 20,000 items
	cluster C	14,178 > 8334 items	9794 < 12,500 items	9182 < 20,000 items
ECLARANS	cluster A	1977 < 8334 items	17,884 > 12,500 items	35,223 > 20,000 items
	cluster B	3529 < 8334 items	10,741 < 12,500 items	15,602 < 20,000 items
	cluster C	19,493 > 8334 items	8874 < 12,500 items	9174 < 20,000 items
KSOM	cluster A	2977 < 8334 items	16,884 > 12,500 items	34,223 > 20,000 items
	cluster B	4529 < 8334 items	11,741 < 12,500 items	16,602 < 20,000 items
	cluster C	17,494 > 8334 items	9874 < 12,500 items	10,174 < 20,000 items

Fig. 4 Cluster size using KSOM



In KSOM algorithm, the following size of clusters was generated which is shown in Fig. 4. Here, cluster A and cluster B are considered a small cluster which is shown in Table 2.

4.2.2 Case 2

In this experiment, we have used MovieLens-1 M dataset. This dataset contains 37,500 ratings from 548 users and 1454 movies (items) where filler items are 60%. The ratings are ranging from 1 to 5. In this dataset, filler items are 22,500 and target items are 15,000.

In ECLARANS algorithm, the following size of clusters has been generated as shown in Fig. 5. Here, we see that cluster B and cluster C are small-size clusters as in Table 2. This can be seen that user is included in clusters B and C that are considered as attack profiles.

In PAM algorithm, the following size of clusters was generated as shown in Fig. 6. As per the definition of small cluster, here cluster B and cluster C are considered as small clusters from Table 2.

In KSOM algorithm, the following size of clusters was generated as shown in Fig. 7. Here, cluster B and cluster C are considered as small clusters as in Table 2. This can be seen that user is included in clusters B and C that are considered as attack profiles.

Fig. 5 Cluster size using ECLARANS

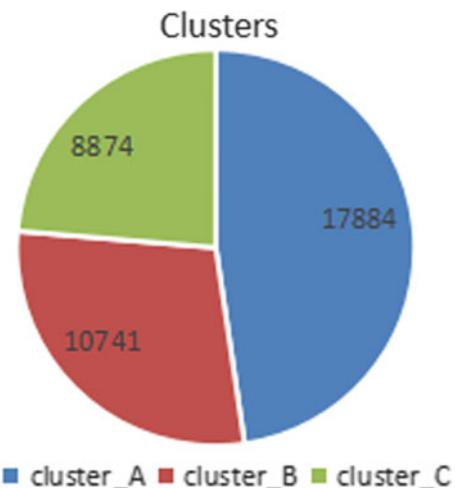
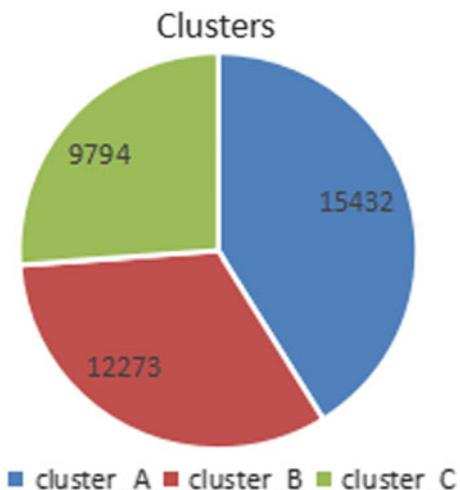


Fig. 6 Cluster size for 60% filler items using PAM



4.2.3 Case 3

In this experiment, we have used MovieLens 1 M dataset. This dataset contains 60,000 ratings from 640 users and 1556 movies (items) where filler items are 70%. The ratings are ranging from 1 to 5. In this dataset, filler items are 42,000 and target items are 18,000.

In *ECLARANS* algorithm, the following size of clusters has been generated as shown in Fig. 8. Here, we see that cluster B and cluster C are small-size clusters from Table 2. This can be seen that user in the dataset is included in clusters B and C that are considered as attack profiles.

Fig. 7 Cluster size using KSOM

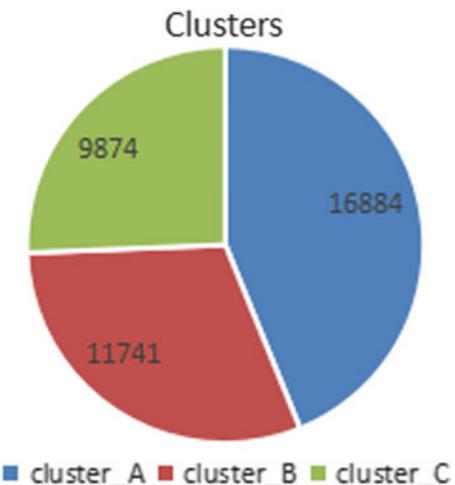
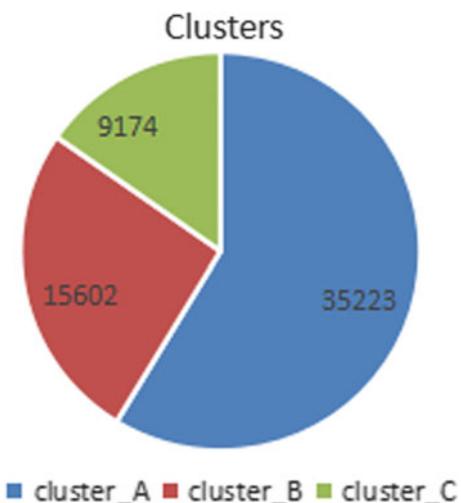


Fig. 8 Cluster size using ECLARANS



In PAM algorithm, the following size of clusters was generated as shown in Fig. 9. As per the definition of small cluster, here, cluster B and cluster C are considered as small clusters that can be seen in Table 2.

In KSOM algorithm, the following size of clusters has been generated as shown in Fig. 10. Here, we see that cluster B and cluster C are small-size clusters as in Table 2. This can be seen that user in the dataset is included in clusters B and C that are considered as attack profiles.

Fig. 9 Cluster size using PAM

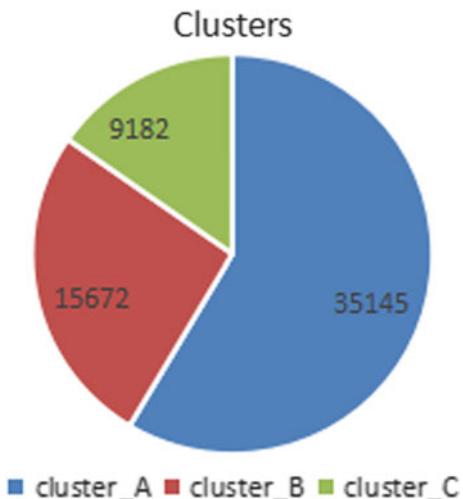
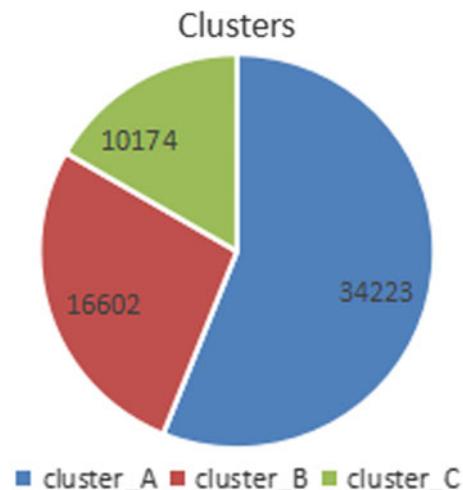


Fig. 10 Cluster size using KSOM



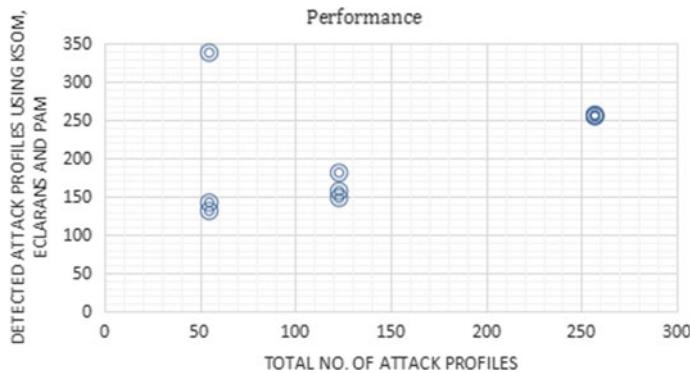
4.2.4 Performance

One of the measures is the outlier accuracy given in Eq. 4 that we have used here is to check the performance accuracy between KSOM, ECLARANS, and PAM algorithms that tells that the profile is not under the region of other group profiles, namely, genuine and fake. The results are shown in Table 3.

$$\text{Outlier Accuracy} = \frac{\text{Total no. of Attack Profiles} * 100}{\text{No. of Attack Profiles detected using KSOM or ECLARANS or PAM}} \quad (4)$$

Table 3 Profile injection attacks detection using KSOM, ECLARANS, and PAM

Filler Items (%)	40%	60%	70%
# Users	547	548	640
# Attack profiles	55	123	257
# Attack profiles detected through PAM	338	181	257
# Attack profiles detected through ECLARANS	142	157	255
# Attack profiles detected through KSOM	131	148	255
PAM algorithm performance to detect attack profiles	16%	68%	100%
ECLARANS algorithm performance to detect attack profiles	38.73%	78.34%	99.22%
KSOM algorithm performance to detect attack profiles	41.98%	83.10%	99.22%

**Fig. 11** Detection of attack profiles in 40, 60, and 70% filler items by KSOM, ECLARANS, and PAM

In the existence of ECLARANS and PAM clustering algorithm, it has been noticed that KSOM algorithm for the detection of shilling attack on different filler items is more, which is shown in Fig. 11.

5 Conclusion

Fake profiles and genuine profiles of filler items have been detected using KSOM algorithm, and outlier accuracy performance of KSOM is compared with already existing ECLARANS and PAM unsupervised clustering techniques. In the future endeavor, there will be a focus on some other clustering/classification techniques for the optimization of outliers' detection.

Acknowledgements I give my warm thanks to Dr. Veer Sain Dixit as research supervisor for his idea about SOM and encouragement of writing this paper. Without his valuable support, this work is not possible toward enhancement in the knowledge about profile injection attack detection.

References

1. Aggarwal, C.C.: An Introduction to recommender systems (The Textbook), vol. XXI, 1st edn., 498p. Springer International Publishing, Cham (2016). ISBN: 978-3-319-29657-9
2. Chakraborty, P.: A scalable collaborative filtering based recommender system using incremental clustering. In: IEEE International Advance Computing Conference, Patiala, India, 6–7 March (2009)
3. Burke, R., Mobasher, B., Williams, C., Bhaumik, R.: Detecting profile injection attacks in collaborative recommender systems. In: Proceedings of the 8th IEEE International Conference on E-Commerce Technology and the 3rd IEEE International Conference on Enterprise Computing, E-Commerce and E-Services (2006)
4. Aghili, G., Shahri, M., Khadivi, S., Morid, M.A.: Using genre interest of users to detect profile injection attacks in movie recommender systems. In: Proceeding of IEEE International Conference on Machine Learning and Applications (2011)
5. Bryan, K., O'Mahony, M., Cunningham, P.: Unsupervised retrieval of attack profiles in collaborative recommender systems. Technical Report UCD-CSI2008-03, University College Dublin (2008)
6. Dhimmar, J.H., Chauhan, R.: An accuracy improvement of detection of profile injection attacks in recommender systems using outlier analysis. Int. J. Comput. Appl. (0975–8887) **122**(10), (2015)
7. Chakraborty, P., Karforma, S.: Detection of profile-injection attacks in recommender systems using outlier analysis. In: International Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA), Procedia Technology, vol. 10 (2013)
8. Toor, A.K., Singh, A.: Analysis of clustering algorithm based on number of clusters, error rate, computation time and map topology on large data set. IJETTCS **2**(6), 94–98 (2013)
9. Sinwar, D., Kumar, S., Dr.: Study of different clustering approaches for outlier detection. IJCSC **4**(2), 51–54 (2013)
10. Hurley, N., Cheng, Z., Zhang, M.: Statistical Attack RecSys'09, 23–25 October 2009. ACM, New York (2009). 978-1-60558-435- 5/09/10
11. Bilge, A., Ozdemir, Z., Polat, H.: A novel shilling attack detection method. In: Information Technology and Quantitative Management, ITQM 2014 (Proc. Comput. Sci. **2014**, 165–174) (2014). <https://doi.org/10.1016/j.procs.2014.05.257>
12. Dan, J., Jianlinet, Q., Xiang, G., Li, C., Peng, H.: a synthesized data mining algorithm based on clustering and decision tree. In: 10th IEEE International Conference on Computer and Information Technology (CIT 2010)
13. Ai, H., Li, W.: K-means initial clustering center optimal algorithm based on estimating density and refining initial. IEEE (2012)
14. Hurley, N., Cheng, Z., Zhang, M.: Statistical attack detection. In: Proceedings of the Third ACM Conference on Recommender Systems, pp. 149–156. ACM (2009)
15. Kriegel, H.-P., Schubert, M., Zimek, A.: Angle-based outlier detection in high dimensional data. In: Proceeding of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, NY, USA, pp. 444–452 (2008)
16. Al-Zoubi, M.B.: An effective clustering based approach for outlier detection. Eur. J. Sci. Res. **28**(2), 310–316 (2009)
17. Vesanto, J., Alhoniemi, E.: Clustering of the self organizing map. IEEE Trans. Neural Netw. **11**(3), 586–600 (2000)
18. Mishra, M., Behera, H.S.: Kohonen self organizing map with modified K-means clustering for high dimensional data set. Int. J. Appl. Inf. Syst. **2**(3), 34–39 (2012)
19. WEKA 3. <https://www.cs.waikato.ac.nz/ml/weka/>. Accessed 13 Dec 2017
20. MATLAB 9.0. <https://in.mathworks.com/>. Accessed on 13 Jan 2018

Fruition of CPS and IoT in Context of Industry 4.0



Mishra Devesh, Agrawal Krishna Kant, Yadav Ram Suchit, Pande Tanuja and Shukla Narendra Kumar

Abstract The application of sensor networks and Radio Frequency Identification (RFID) is not new to the industries. Recent advancements in communication technologies with more expertise in networking algorithms paved the updation of the Internet toward the foundation of Internet of Things (IoT). This is an interconnection designed for seamless communication between anyone to anywhere from any time to anything with any network for any service that has potential to turn up the whole smart sensor interconnected network into a Cyber-Physical System (CPS). Furthermore, by applying information analytics, interacting machines will complete the task more proficiently, collaboratively, and sturdily. This next generation trend of industry is named as Industry 4.0. This article itemizes an all-encompassing set of technical awareness about the definite concept of IoT, CPS, and Industry 4.0. This paper contributes to summarize the current state-of-the-art technology for interconnected machine operation in industries.

Keywords Wireless sensor network (WSN) · Internet of things (IoT) · Cyber-physical system (CPS) · RFID · Self-configure

1 Introduction

The development of Cyber-Physical System (CPS) gets realized due to the contemporary advances in technology within which data collected from all sensing units is analyzed in all perspectives and get coordinated between the physical industry floor and cyberspace [1]. CPS is an array of multidisciplinary systems that incorporate entrenched computer and digital sensor system technologies into the physical world. Numerous digital sensing and computing features following the modern complex

M. Devesh (✉) · Y. R. Suchit · P. Tanuja · S. N. Kumar

Department of Electronics & Communication, University of Allahabad, Allahabad, India
e-mail: deveshbbs@gmail.com

A. K. Kant

Department of Computer Science & Engineering, Delhi Technical Campus, Greater Noida, India

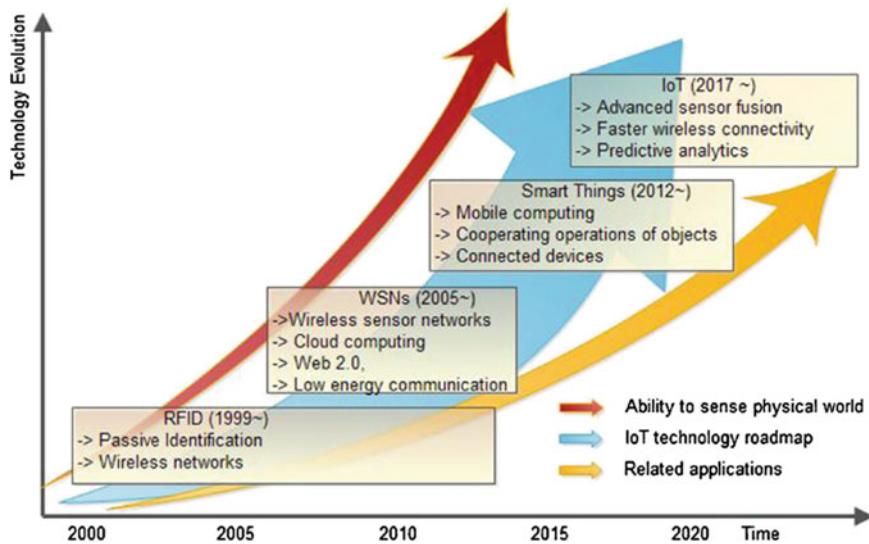


Fig. 1 Growth of IoT [1, 2]

procedure for application necessitate latest approaches which integrate cyber world with a physical system for analysis of terrific data and to establish a communication from man to man, man to machine, and machine to machine at all levels.

With the advancement in smart sensors, intelligent computing, and low energy communication unit forming a network of a large number of things crafted as the Internet of Things (IoT). Things in IoT communicate, exchange, and analyze information autonomously which offers interoperability, trustworthiness, and accuracy of the system. In the year 1999, the concept of IoT got initiated in which Radio Frequency Identification (RFID) technology was referred to as IoT as uniquely distinguishable objects for information exchange. By the end of the year 2013, IoT had become a superset of connected devices which were recognizable by near field communication. The “Internet” and “Things” interconnection of worldwide devices is based on intelligent information exchange with processing, sensing, and analysis in the background. With communication between physical and virtual “things” having intelligent interfaces forming a network. As shown in Fig. 1, IoT originated from RFID and the growing sensor technologies elevated the IoT to autonomous control. This extended the IoT into a new cluster of technologies like sensor networks, cloud computing, barcodes, NFC, etc. Accumulation of all these has supported IoT to get into the next generation of computing, communication, and control where physical things could be accessed through the Internet making CPS a universal set of these interconnected, interoperable, intelligent smart sensing devices to enhance the efficiency, accuracy, quality, and reliability of the system [2].

2 IOT Paradigm and Evolution

The computers and IoT share lots of common blocks including a bundle of technologies. Significantly IoT comprises of lots of different technologies with dissimilar characteristics as per the request of the system. IoT is another step of progression of our existing internet system. The gradual development of network from digital computers from 1960 to the introduction of TCP/IP in 1980 and later www propelled the popularity of the Internet in 1991 and crafted a term “Web of Things” as shown in Fig. 2. In the next generation due to the connection of mobile devices with the internet, scribes get connected through the Internet forming social networking. The fifth stage of development of the Internet nevertheless called as the IoT has all the physical, computable, sensor-based digital objects all-around that get connected to each other with or without any human intervention with the use of the Internet [3].

3 Components of IoT

The functionality of IoT is divided into six categories as identification, sensing, communication, computation, services, and semantics. Different technologies are associated at each level of classification. The systematic combination of technology from each step is responsible for the interconnected network of things via the Internet known as IoT [5] (Fig. 3).

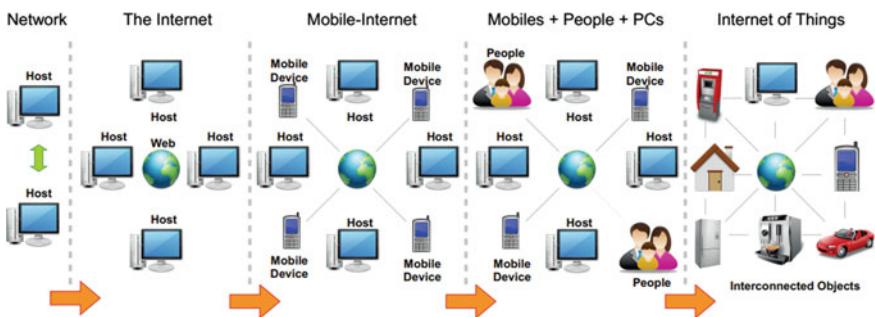


Fig. 2 Phases of IoT evolution [3, 4]



Fig. 3 Foundations of IoT [5]

3.1 Identification

Addressing and identification of device are vital components in IoT as enormous components communicate and exchange information in an IoT network. For identification of devices Electronic Product Code (EPC), Ubiquitous Codes (uCode), and many more techniques are available. Identification of object in IOT refers to its ID in the network while the location is device address within the IoT network either as source or sink. Object identification and addressing are dissimilar to each other meanwhile identification are universally matchless, therefore addressing succors to exclusively recognize things.

3.2 Sensing

Smart sensors are used for gathering information within the network about related parameters from the surrounding environment and the collected data is analyzed and transmitted to the cloud, which is a database of clear location for essential services. With smart sensors and actuators, IoT offers smart hubs facilitating users to control thousands of devices and appliances autonomously.

3.3 Communication

The communication in IoT between dissimilar, heterogeneous objects noted as things provides medium to deliver and exchange information in the network of thousands of smart devices. Things in IoT should run with low power and lossless communication protocol. Few major communication protocols used in IoT are WiFi (IEEE802.11 a/b/g/n), Bluetooth (IEEE802.15.1), Zigbee (IEEE802.15.4), Long-Term Evolution (LTE), RFID, NFC, and ultrawide bandwidth (UWB).

3.4 Computation

Microcontrollers, microprocessors, Field Programmable Gate Arrays (FPGA), and Application Specific Instruction set Computing (ASIC) provide computational strength to the nodes of IOT within the network for processing and analyzing accumulated sensed data enhancing self-configuring with autonomous decision-making the ability of the network. Various hardware platforms are used for the development of IoT such as different versions of Arduino , AVR, ARM, PIC, Raspberry

PI, Orange PI, BeagleBone, WiSense, and T-Mote Sky, etc. Moreover, software platforms are used to offer an easy user interface to IoT nodes like Real-Time Operating Systems (RTOS), TinyOS, LiteOS, and Riot OS provides OS intended for IoT atmosphere [6].

3.5 Services

Services in IoT can be categorized into Identity-related Services, Information Aggregation Services, Collaborative-Aware Services, and Ubiquitous Services. The basic and important service that is used in IoT is Identity-related services which brings real-world objects to the virtual world. Information Aggregation Services encapsulate the gathered raw sensory depths that must be conveyed to the IoT application. Collaborative-Aware Services act on top of the previous one and use the gathered data for making a decision and reacting consequently. Smart home and smart grid comes into the category of information aggregation services, while industrial automation falls in collaborative aware services. Anyone, anytime, and anywhere service of Collaborative-Aware Services is delivered by Ubiquitous Services [7].

3.6 Semantics

The information retrieval section of heterogeneous, smart sensor IoT things comes in this category of semantics in IoT. Information retrieval includes discovering and using resources and modeling information. Also, it consists of identifying and analyzing data to make logic of the right choice to provide the precise facility. Thus, semantic represents the brain of the IoT by sending demands to the right resource.

4 Relationship Between Sensor Network and IoT

There is an acute relationship between wireless sensor network and IoT. Sensor networks can occur without the IoT. Conversely, the IoT cannot be present without sensor network, because the majority of hardware for sensing, communication, and computation belongs to mote moreover structure support, including access to sensors and actuators, also belongs to wireless sensor networks. Sensor networks are part of IoT while IoT is not part of the sensor network [8].

As shown in Fig. 4 multiple numbers of wireless sensor nodes called as motes are connected in wireless sensor network relaying information from the source node to the sink node in a multi-hop manner. In traditional fashion homogenous motes are used but nowadays for multimodal sensing performance, heterogeneous sensor motes may be applied which exchange information and communicate between themselves

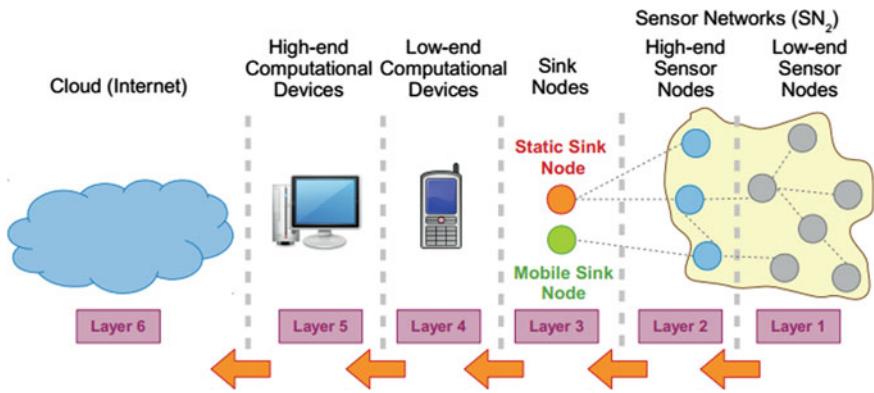


Fig. 4 Stratums of sensor network with IoT [4]

via different technologies and protocols of the Internet. The IoT follows a six-layer architecture data transfer between nodes to the cloud consisting of mobiles device and computers in between. There are many technologies available to craft a wireless network of motes such as wireless personal area network (WPAN), wireless local area network (WLAN-Wi-Fi), wireless metropolitan area network (WMAN-WiMAX), wireless wide area network (WWAN-2G and 3G), and satellite network (GPS). Sensor networks are used in a restricted domain with specific purpose and parameter of monitoring like in case of structural health monitoring of pipelines, environment monitoring, agriculture, and healthcare surveillance. Therefore, motes of the sensor network use two types of protocol for data communication as non IP-based communication protocol like Zigbee and IP-based protocol comprises of IPv6 [4].

5 Data Flow Structure from Sensor Network to IoT

The roots of CPS and IoT are embedded in the soil of wireless sensor network because no CPS and IoT can be realized without the application of wirelessly interconnected smart sensor nodes with the application of different technologies at each stage and finally in the whole sum we obtain Industry 4.0 as the fruits of this system as shown in Fig. 5.

Applying CPS in industries of the modern era is classified into three stages as component, machine, and production. At the component stage, the data from the sensors gathering information about the critical components get converted into information after analyzing the parameters being monitored in a self-optimizing, self-protect manner. At the next stage, all the machine tools provide data to the controller for aggregation of multimodal information gathered at the component level. And at the production level, the machines couple with CPS provide self-maintainability and

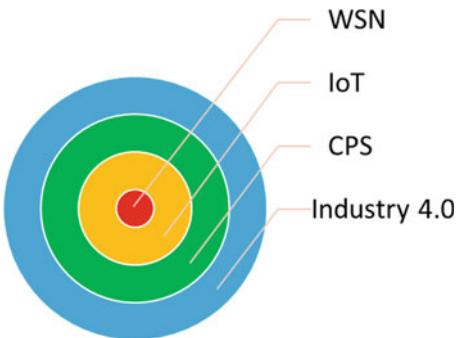


Fig. 5 Relationship between wsn and IoT

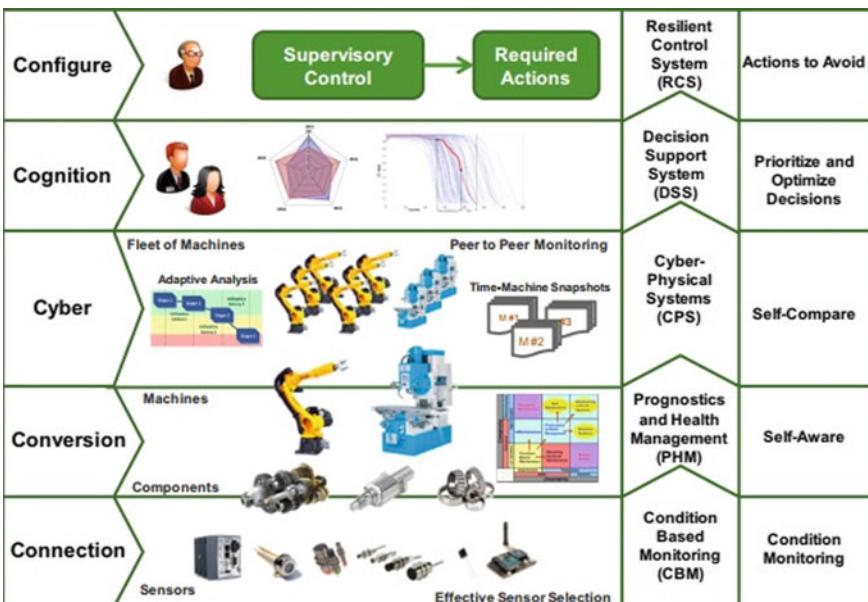


Fig. 6 Technology and use at each level [9]

self-configurability to the machine tools inside the industry collecting all the gathered data at each level of operation [9] (Fig. 6).

Obtaining data from critical components of machine, tools, or infrastructure in the industry through smart sensor interconnected network of the Internet is prerequisite part of building CPS in any industry. Two opportunities are available to perform this action. First is to gather information directly from the sensor that are provided to the controller for further processing and the second is having smart sensing capability making the device equipped with a microcontroller to analyze the information gathered from sensing unit.

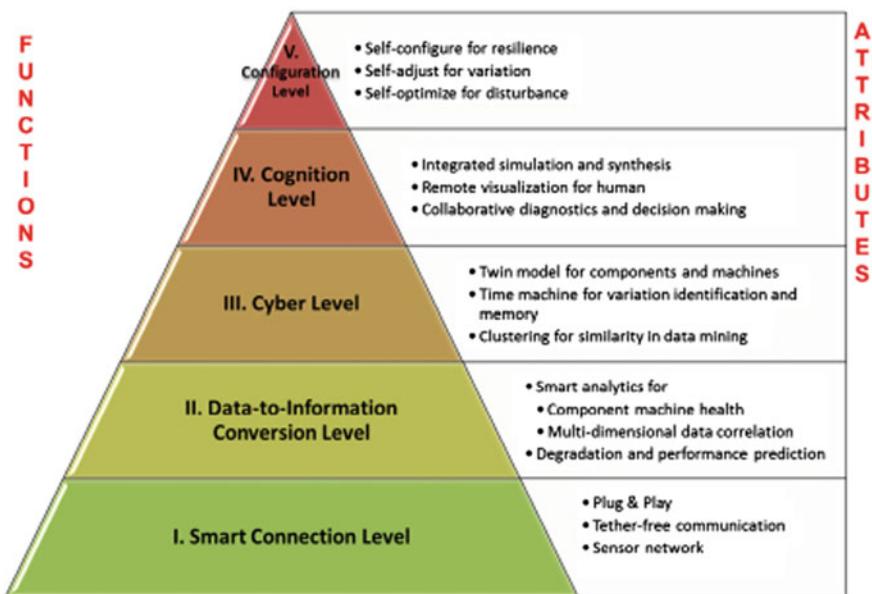


Fig. 7 Pyramid of 5C design [9]

The second step is the conversion of data to useful information, in technical terms it is an information retrieval stage. Significant information is extracted from all the data gathered in multimodal pattern from all the interconnected homogeneous or heterogeneous motes within the network boundaries of the Internet. Performance and the loss in the reliability of the machine can easily be determined at this stage. In combination with the previous layer, this stratum provides self-awareness to the monitored machine [5] (Fig. 7).

This is a central stage of architecture and acts as a pivot of information received in from heterogeneous sources with multimodal parameters. This stage sometimes uses big data algorithms to extract meaningful information about machine status. This analysis at this phase offers the self-comparison ability to the system. This delivers competence to equate present machine performance with the previous one.

The intervention of human beings for decision-making, diagnosis, visualization, and synthesis occurs at this level in collaboration of the previous three levels. Relative data with specific machine status is presented, and decision on the precedence of responsibilities can be made.

The final step is feedback from cyberspace to the physical world and it has a regulatory control to make self-configurable for resilience, self-adjust for variation, and self-optimize for a disturbance [10].

6 Conclusion

In the present day scenario the process industry uses sensor, microcontroller equipped machines, and monitoring at production phase, but currently all the attributes are working with isolation from other stages, while in industry 4.0 the sensors are embedded in microcontroller crafting smart sensors equipped with low power communication device having the capability of forming a networked structure via the Internet. This mechanism fuels up the traditional industry with skills similar to self-aware, self-predict, self-compare, self-configure, self-maintain, and self-organize.

When cyber set-up is installed, machines get the index into the network and interchange data through cyber interfaces. A system could be developed for the pathway the changes of a machine present have supplementary information about the previous stages of the machine or devices connected within the network for regular monitoring of present output obtained from the previous one.

References

1. Li, S., Da Xu, L., Zhao, S.: The internet of things: a survey, 26 April 2014. Springer Science + Business Media, New York (2014)
2. Torgul, B., Sağbanşua, L., Balo, F.: Internet of things: a survey. *Int. J. Appl. Math. Electron. Comput.* (2016). ISSN: 2147-8228
3. Perera, C., Liu C.H., Jayawardena, S., Chen, M.: Context-aware computing in the internet of things: a survey on internet of things from industrial market perspective. *IEEE ACCESS* (2016)
4. Perera, C., Zaslavsky, A., Christen, P., Georgakopoulos, D.: Context-aware computing for the internet of things: a survey. *IEEE Commun. Surv. Tutor.* **16**(1), First Quarter 2014
5. Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., Ayyash, M.: Internet of things: a survey on enabling technologies, protocols, and applications. *IEEE Commun. Surv. Tutor.* **17**(4) (2015)
6. Da Xu, L., He, W., Li, S.: Internet of things in industries: a survey. *IEEE Trans. Ind. Inf.* **10**(4) (2014)
7. Gunes, V., Peter, S., Givargis, T., Vahid, F.: A survey on concepts, applications, and challenges in cyber-physical systems. *KSII Trans. Internet Inf. Syst.* **8**(12) (2014). Copyright © 2014 KSII
8. Khaitan, S.K., McCalley, J.D.: Design techniques and applications of cyberphysical systems: a survey. *IEEE Syst. J.* **9**(2) (2015)
9. Lee, J., Bagheri, B., Kao, H.A.: A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. Society of Manufacturing Engineers (SME). Elsevier Ltd. All rights reserved (2014)
10. Jia, D., Lu, K., Wang, J., Zhang, X., Shen, X.: A survey on platoon-based vehicular cyber-physical systems. *IEEE Commun. Surv. Tutor.* **18**(1) (2016)

Leveraging Biological Dragonfly Scheme for URLLC in Industrial Wireless Network



Sanjay Bhardwaj, Muhammad Rusyadi Ramli and Dong-Seong Kim

Abstract The dragonflies have behaviours, i.e. the formation of the swarms as the static and the dynamic. The formation is the representation of the optimization, not only in terms of the identification of the food source but also a distraction from the enemy. So, these behaviours inspire a novel technique and algorithm for an ultra reliable low latency communication in an industrial wireless network which is proposed as URLLC-DF. The convergence of the URLLC-DF towards a global optimum and in comparison with other bio-inspired evolutionary algorithm is much more efficient in a better way.

Keywords Dragonfly · Ultra reliable low latency communication · Industrial wireless network · Convergence · Optimization

1 Introduction

The plants as well as animals are full of behaviour and patterns that are socially linked with each other. This social behaviour in a way is the means of performing the different tasks, which otherwise might be improbable to do individually. The ultimate goal of these interactions and collective behaviour is survival, defend, hunt, navigation and foraging. For mobile ad hoc networks (MANETs), [1], propose a protocol which is dependent and having knowledge of the geographical area, which

S. Bhardwaj · M. R. Ramli

Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi, South Korea

e-mail: sanjay.b1976@gmail.com

M. R. Ramli

e-mail: muhammadrusyadi@kumoh.ac.kr

D.-S. Kim (✉)

IEEE, Gumi, South Korea

e-mail: dskim@kumoh.ac.kr

is a hybrid routing protocol, called as the zone routing protocol (ZRP), in which the latency is improved with only marginal fall in the packet delivery ratio.

Wolves are considered ferocious predators, but even they hunt in packs, which shows their social and organizational skill [2, 3], proposed an optimization algorithm, which is also inspired from the wolves, called as the wolf search algorithm (WSA), imitating their survival, hunting skills, and avoiding of the predators.

Ants are considered at the highest level of the social intelligence, and this behaviour which is called as the swarm intelligence (SI) was first described as ant colony optimization (ACO) [4]. Another SI paradigm called as particle swarm optimization (PSO) algorithm which replicates navigational characteristics of the flock of birds [5], and is based on the interaction of the birds among themselves. Another new and prevalent SI-based algorithm, the artificial bee colony (ABC), which simulates the social foraging nectar behaviour of honey bees, [6].

Dragonfly optimization (DFO), which is inspired from the formation of the swarms by the dragonflies when they hunt, search for the food and avoid the predators, have been used for solving combinational optimization problems, [7].

The remaining of the paper is organized as follows: Sect. 2 presents the dragonfly behaviour and ultra reliable low latency communication (URLLC) characteristics, which acts as the inspiration and motivation of the paper, Sect. 3 explains about the DFO and its subsection explains about initialization, static and dynamic swarms, velocity upgradation, mutation and information sharing, the proposed algorithm, Ultra Reliable Low Latency-Dragonfly Algorithm(URLLC-DF), is discussed in Sect. 4, the results are discussed in Sect. 5 and Sect. 6 discusses the conclusion and future work.

This instruction file for Word users (there is a separate instruction file for LaTeX users) may be used as a template. Kindly send the final and checked Word and PDF files of your paper to the Contact Volume Editor. This is usually one of the organizers of the conference. You should make sure that the Word and the PDF files are identical and correct and that only one version of your paper is sent. It is not possible to update files at a later stage. Please note that we do not need the printed paper.

We would like to draw your attention to the fact that it is not possible to modify a paper in any way, once it has been published. This applies to both the printed book and the online version of the publication. Every detail, including the order of the names of the authors, should be checked before the paper is sent to the Volume Editors.

2 Dragonfly Behaviour and URLLC Characteristics

Dragonflies are the smallest of the predators but have the highest kill ratio among all. This high kill ratio can be attributed because of the rare swarming behaviour: Static (feeding) Swarm and Dynamic (migratory) swarm. The static and dynamic swarms have their own characteristics which are useful in the survivability of the dragonflies. The formation of the static swarm helps the dragonfly to hunt for the

small preys such as butterflies and mosquitoes [8] whereas in case, the dragonflies want to migrate over a long distance they tend to form dynamic swarms, [9].

The generation of the swarms helps them in the search of the food and keeps them safe from the predators as well i.e. the formation of the swarms is required for their survivability. The change from the static (small) to the dynamic (big) swarm also shows the self-organizational skills they have and in perfect unison they are able to scale. Thus, the dragonflies have three basic qualities: Scalability, Self-Organization and Survivability, which are also required for the basic communication system. These three characteristics can be further extended in relation to the three primitive principles of the swarms: Separation, Alignment, Cohesion [10].

Thus, the motivation for the use of the dragonfly, is derived by drawing an analogy between the behaviour of the dragonflies and the URLLC, which is the separation (static swarm behaviour) improves the reliability and alignment along with cohesion (dynamic swarm behaviour) gives a low latency wireless network. Thus, the algorithm which is called as URLLC-DF attains the local best and the global best solutions that satisfies the basic requirements of the communication for it to be called so.

3 Dragonfly Optimization

The main characteristic behaviour of the dragonfly are separation, alignment and cohesion, which defines the formation of the static and dynamic swarms as alignment should be high to maintain an appropriate separation even if cohesion is low in dynamic swarm, but for static swarm high cohesion is required to attack the preys even if the alignment is low. Individual dragonflies with the help of the additional parameters make optimum use of search space use and as a result of which avoid premature convergence, [12]. So, weights are assigned to the dragonflies accordingly whether they are in the dynamic or static swarm. So, each dragonfly can be defined by the parameters:

$$df_i = df_i^1, df_i^2, \dots df_i^{dv}, s_i, a_i, c_i, f_i, e_i, w_i, \quad (1)$$

where, s_i is separation, a_i is alignment, c_i is cohesion, f_i is food, e_i is the enemy of the i th dragonfly.

3.1 Initialization, Static and Dynamic Swarms Formation

Before the iterative process is started the population members of the dragonfly are initialized, Eq. 2, so as to be in the required search space and uniformly distributed. With all the initial positions fixed along with the uniform distribution, formation of

the swarms takes place which can be static as well as dynamic. This formation of swarms helps them in the search of the food and defends them from the predators.

$$df_i^j = [df_i^1, df_i^2, df_i^3, \dots df_i^{dv}], \quad (2)$$

where, df_i^j indicates j th position of the i th dragonfly and the problem decision variables are represented by dv .

3.2 Velocity, Position Update, Mutation and Information Sharing

With the continuous movement of the dragonflies, their velocity and position are updated at any moment of time ($t + 1$) Eq. 3. This updated position and the velocity vector of the individual as well that of the swarms helps in maintaining perfect synchronization in terms of the information sharing of the food sources as well that of the predators.

$$df_j(t_j + 1) = (s_j \times S_{df_j} + a_j \times A_{df_j} + c_j \times C_{df_j} + f_j \times F_{df_j} + e_j \times E_{df_j} + w_j \times V_{df_j}(t_j)), \quad (3)$$

where, s_j , a_j , c_j , f_j and e_j are the factors which govern the formation of the swarms, Eq. 1, w_j is weight factor, t_j is the counter, for the j th dragonfly, for all $j = i$.

For mutation, Eq. 4, captures the real essence of the swarm formations, making sure that the swarm M_{si} have almost the same number of the dragonflies and it is further extended by random operator *rand*.

$$M_{si}(t + 1) = M_{si}(t) + rand \times (rand - M_{si}(t)). \quad (4)$$

4 System Model

For a communication system to be considered as the URLLC, it should possess atleast these three characteristics. The separation (avoidance of the collisions) which also increases the survival rate of the dragonflies as they can focus on the ‘food source’ i.e. the ‘target’ and prevent them from the ‘enemies’ i.e. the ‘noise source’ gives a highly reliable communication system and alignment (velocity/energy matching) along with cohesion (teamwork) i.e. ‘better coordination, exchange of information, energy conservation’ results in a low latency wireless network. In particular, with respect to the URLLC, in addition, it should have high reliability, high Signal to Noise Ratio (SNR) and least amount of delay, [11].

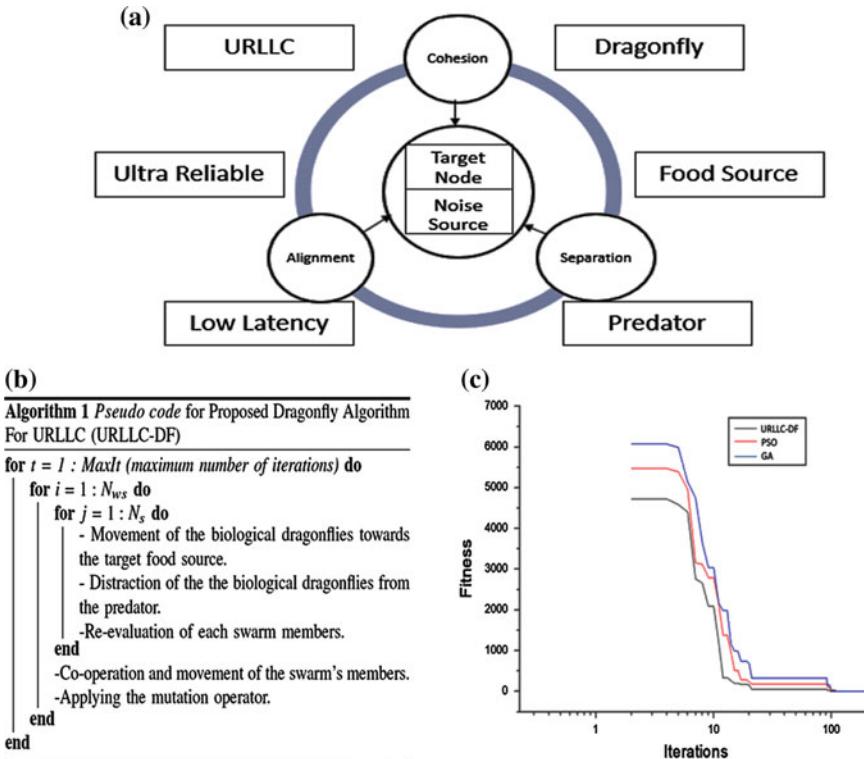


Fig. 1 **a** System model. **b** Pseudo-code for URLLC-DF. **b** Convergence curve

Thus, URLLC, has two basic things associated with it: ultra reliability and low latency and on the similar lines, dragonfly also has two identification characteristics: food source and predator, along with the basic behaviour of cohesion, alignment and separation. Thus, in the system model, the dragonflies' characteristics as well the basic behaviour supports the search and identification of the target node as well as that of the noise source. Thus, in totality, the dragonfly can be modelled into the URLLC, Fig. 1a.

5 Proposed Algorithm (URLLC-DF)

URLLC, is a multi-objective optimization protocol, which can be modelled in along with the dragonfly and collectively called as URLLC-DF, Fig. 1b. The convergence towards the food source—‘target node’ and the distraction from the predator—‘Noise source’, Fig. 1c. During this process of the convergence, the information sharing between the swarms takes place, which is also regularly updated and the mutation

operator thus making sure that the dragonflies in the swarms are able to move freely from one swarm to another thus avoiding premature convergence. The robustness, ability to converge, along with the identification of the target source in the shortest possible time is further explored, discussed and strengthened that the URLLC-DF, Fig. 1b, is able to deliver on the proposed lines in Sect. 6.

6 Result and Discussion

Figure 1c, shows the convergence curves: number of dragonflies is taken as 35 and the maximum number of iterations is 200. In this convergence curve, a comparison is made in between the proposed algorithm URLLC-DF algorithms: genetic algorithm (GA) and particle swarm optimization (PSO). From the trajectory of the fitness values in Fig. 1c, of the dragonflies, it is clearly observed, the decrease over the course of iterations. From which it is clearly conceptualized that the URLLC-DF has a better convergence rate as compared to the other two, also as the number of the iterations go on increasing the convergence gets accelerated, which is attributed to more focus on the exploitation in comparison to the exploration behaviour of the dragonfly. This behaviour guarantees that the algorithm eventually converges to a point and search space is optimized along with the optimum value [12].

7 Conclusion and Future Work

The proposed algorithm, URLLC-DF, performed better as the convergence curves from the first to the last iteration is able to converge at a much faster rate than PSO and GA. So, in terms of the Communication System, identification of the target source and the noise source: gives a highly reliable medium and a fast convergence gives out the low latency communication system. So, as per the findings of this comprehensive study, shows that the algorithm URLLC-DF outperforms the current and well-known algorithms in the literature and can be evidently used for URLLC in the industrial wireless networks. As the future work, a hybrid approach with another evolutionary algorithm can be explored.

Acknowledgements This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the Global IT Talent support programme (IITP-2017-0-01811) supervised by the IITP (Institute for Information and Communication Technology Promotion) and National Research Foundation of Korea (NRF) through the Advanced Research Project (NO. NRF-2017R1A2B4009900).

References

1. Pham, T.M.T., Nguyen, T. T., Kim, D.S.: Geographical awareness hybrid routing protocol in mobile ad hoc networks. *IEEE Trans. Ind. Inform.* **1**(8), 61–68 (2017)
2. Saremi, S., Mirjalili, S.Z., Mirjalili, S.M.: Evolutionary population dynamics and grey wolf optimizer. *Neural Comput. Appl.* **5**(26), 41257–1263 (2015)
3. Fong, S., Deb, S., Yang, X.S.: A heuristic optimization method inspired by wolf preying behavior. *Neural Comput. Appl.* **7**(26), 1725–1738 (2015)
4. Dorigo, M., Stutzle, T.: *Handbook of Metaheuristics*. Springer, Switzerland (2003)
5. Eberhart, R., Kennedy, J.: A new optimizer using particle swarm theory. *IEEE* (1995)
6. Karaboga, D.: Technical report, Erciyes University (2005)
7. Mirjalili, S.: Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. *Neural Comput. Appl.* **4**(27), 1053–1073 (2016)
8. Wikelski, M., Moskowitz, D., Adelman, J.S., Cochran, J., Wilcove, D.S., May, M.L.: Simple rules guide dragonfly migration. *Am. Midl. Nat.* **3**(2), 325–329 (2006)
9. Russell, R.W., May, M.L., Soltesz, K.L., Fitzpatrick, J.W.: Massive swarm migrations of dragonflies (Odonata) in eastern North America. *Am. Midl. Nat.* **2**(5), 325–342 (1998)
10. Reynolds, C.W.: Flocks, herds and schools: a distributed behavioural model. *ACM SIGGRAPH Comput. Graph.* **4**(21), 25–34 (1987)
11. Chen, H., Abbas, R., Cheng, P., Shirvanimoghaddam, M., Hardjawana, W., Bao, W., Li, Y., Vucetic, B.: Ultra-reliable low latency cellular networks: use cases, challenges and approaches (2017). [arXiv:1709.00560](https://arxiv.org/abs/1709.00560)
12. Sambandam, R.K., Jayaraman, S.: Self-adaptive dragonfly based optimal thresholding for multilevel segmentation of digital images. *J. King Saud Univ.-Comput. Inf. Sci.* (2016)

Recent Advances in Artifact Removal Techniques for EEG Signal Processing



Amandeep Bisht, Chamandeep Kaur and Preeti Singh

Abstract EEG recordings are frequently contaminated with unavoidable artifacts. Preprocessing in EEG has been a dynamic field of investigation as none of the reported methods can be reviewed as a standard approach for effective artifact removal. This paper presents a broad survey of the existing artifact removal methods. This review is expected to help researchers in improving the existing artifact handling techniques so as to relieve the expert's burden by ensuring efficient analysis.

Keywords EEG · Artifacts · Signal processing

1 Introduction

EEG is used to record the electrical activity within the brain. EEG analysis is generally carried out by trained experts who visually scan the EEG recordings. This task of visual scanning is very time-consuming and depends on one's perception making the analysis complex and unreliable. These artifacts overlap the useful neural information making it hard to distinguish for further applications. In general, there are two main sources of artifacts: internal sources and external sources [1]. Internal artifacts are generated from the subject's physiological activity. It includes electrooculography (EOG) artifacts (due to eye movement, sleep and eye blinking /movement/flattering), cardiac artifacts, electromyography (EMG) artifacts (due to

A. Bisht · C. Kaur · P. Singh (✉)
UIET, Panjab University, Chandigarh, India
e-mail: preeti_singh@pu.ac.in

A. Bisht
e-mail: amandeepbisht@gmail.com

C. Kaur
e-mail: ckchauhan.86@gmail.com

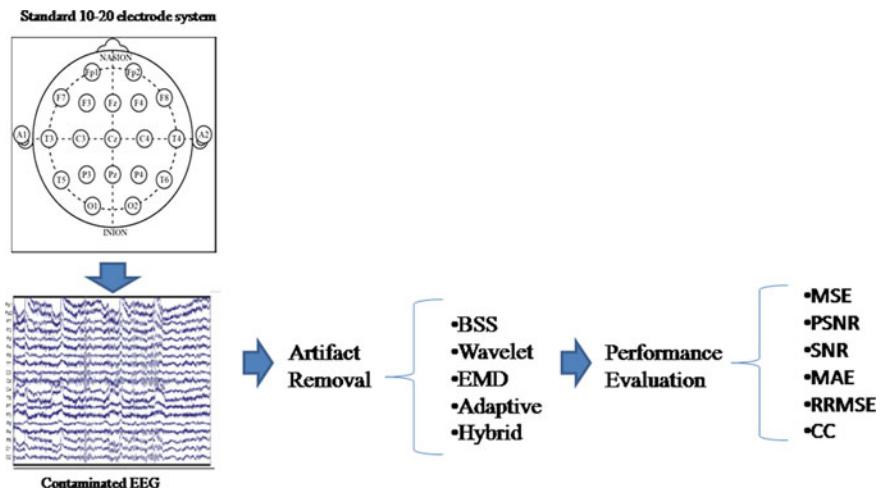


Fig. 1 Steps for EEG processing

muscle contraction, chewing, swallowing, talking, sniffing and clenching) and others (due to skin and respiration activity). Alternatively, external sources arise due to the external environment such as powerline interference (PLI), electrode displacement, pop-up cable noise, poor grounding, wideband interference, and external movements. Some of the artifacts occur at periodic intervals while some at extremely irregular intervals. All these artifacts may be present in a single channel or multiple channels.

Figure 1 presents the workflow for preprocessing mechanism of EEG for different application. For EEG database generation, electrodes are placed on scalp and simultaneously application software and a computer is used to store data. In preprocessing stage, segmentation and artifact removal is done. For performance analysis various qualitative parameters such as mean square error (MSE), signal to noise ratio (SNR), peak signal to noise ratio (PSNR), mean absolute error (MAE), relative root mean square error (RRMSE), and correlation coefficient (CC) are used.

This paper intends to briefly review the existing state of art EEG signal processing via artifact detection as well as removal method implemented on EEG for various applications. Section 2 presents the existing artifact removal methods. Section 3 briefly describes the related work. Lastly, the conclusion has been discussed in Sect. 4.

2 Existing Artifact Removal Techniques

Identification of artifacts is the most important step in preprocessing. In the temporal and spectral domain, use of simple filtering is not sufficient as the spectrum of these artifacts overlaps with EEG signals. In real-time applications, characteristics of artifact and signal must be known for efficient processing. Artifact identification

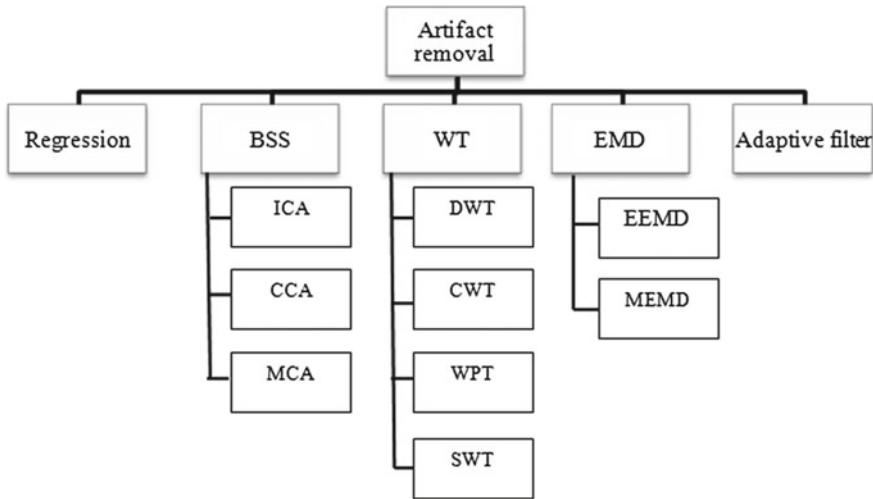


Fig. 2 Flowchart of artifact removal methods

refers to detecting a particular epoch having characteristics similar to the artifact. Figure 2 represents a flowchart of artifact removal methods available in literature.

2.1 Regression Analysis

Regression analysis [2] is a traditional way of recognizing the samples having artifacts and further suppressing those samples not belonging to the system. Nevertheless, major limitation is the requirement of the reference channel. Furthermore, linear regression is not feasible for highly non-stationary signal applications like EEG. In addition, it can process limited artifactual sources, not all categories.

2.2 Blind Source Separation (BSS)

BSS is one of the widely acknowledged methods for artifact suppression which is defined on the basis of unsupervised learning. In this technique, it is presumed that the sources of signals do not correlate with each other i.e. they are statistically independent. Here, it is assumed that the number of observed channels must be greater than or equal to that of signal sources. BSS includes ICA, CCA, and MCA.

(a) Independent component analysis (ICA)

In this BSS-based technique, the estimated sources are expected to be statistically independent [3]. Firstly the data is preprocessed by centering and whitening. Then

optimization of unmixing matrix is done such that it maximizes the non-gaussianity of independent sources.

(b) Canonical correlation analysis (CCA)

It is another method based on BSS which is used for splitting contaminated signals using second-order statistics (SOS). CCA used the first data set as original signal and its time-scaled form as the second set of data. Using SOS it calculates sources that are mutually uncorrelated and maximally auto-correlated and arranges them in ordered form ranging to least auto-correlation from maximum auto-correlation [4].

(c) Morphological component analysis (MCA)

The signal is decomposed into components having diverse morphological properties. Each of these components is characterized sparingly in a dictionary which is over-complete [5]. The effectiveness of this technique is significantly determined by the accessibility of the original artifactual database.

2.3 Wavelet Analysis

Wavelet transform (WT) is a time-frequency based analysis where a signal is decomposed into dilated and translated adaptation of basis functions known as the mother wavelet. There are numerous methods categorized on the basis of wavelet theory, for example, discrete wavelet transform (DWT), continuous wavelet transform (CWT), wavelet packet transform (WPT), and stationary wavelet transform (SWT). Out of all, DWT is the most frequently used method. The procedure of artifact removal is followed by applying the concept of thresholding onto the detailed and approximate coefficients obtained after decomposition by DWT. Then the signal is reconstructed by adding up the newer coefficients [6].

2.4 Empirical Mode Decomposition (EMD)

EMD is completely data-driven whereas wavelet analysis is based upon the selection of mother wavelets (so EMD more flexible and adaptive). The input signal is decomposed into finite amplitude or frequency modulated oscillatory modes called intrinsic mode functions (IMFs). Although EMD provides a complete reconstruction of the data with selective removal of artifactual IMFs it is computationally complex and it is highly sensitive to noise thereby making it a not so good option for online applications [7].

2.5 Adaptive Filtering

In this time variant self-modifying system, a linear filter having a variable transfer function is controlled by adjustable parameters. The feedback from the output is used to adjust the filter parameters so as to minimize error/cost function (difference between desired and observed signal). A reference signal is required to compare filter output and the desired one. An optimization algorithm (such as least mean square or recursive least square) is used to modify these parameters [8].

3 Related Work

Digital filtering is a widely used technique for stationary signals. However, it is not a good option for EEG analysis as it introduces dc offset. For establishing a statistical relationship between signals, regression analysis was introduced. Jung et al. has used regression for removal of channel-specific ocular artifact sources from EEG records [9]. This method results in overcompensation of ocular artifacts and is a very time-consuming method unsuitable for real-time applications. Recently, BSS-based methods have been widely used for ocular artifact (OA) removal. Out of which, ICA-based denoising is efficient in ocular artifact removal but the main drawback of IC-based analysis is computational complexity and the manual selection of artifactual ICs. Moreover, it is applicable to multichannel data only. However, the solution to manual intervention was proposed by using kurtosis, temporal, spatial, and spectral features for automatic ICs detection [3, 10]. Similarly, CCA is capable of segregating muscle artifacts (MA) efficiently as they have a comparatively lower auto-correlation but still the drawback of automation and complexity persists [4]. Likewise, Yong et al. proposed MCA for removal of both ocular artifacts and muscle artifacts. However, this method is appropriate only for those artifacts whose morphology is known [5]. Therefore, the use of BSS techniques as an individual method is not sufficient for artifact removal.

These time domain methods cannot provide any information about frequency domain. Wavelet transform is a time-frequency based method providing longer time duration for information at low frequencies and vice versa for higher frequencies. Khatun et al. in [6] have compared DWT and SWT for different basis functions using either statistical threshold (ST) or universal threshold (UT). The result shows that DWT with ST is suitable for removal of ocular artifacts as ST helps in preserving neural information in EEG. However, DWT-ST had the fastest execution time making it a better option for single-channel applications.

As DWT considers the same basis function for an entire data set that may result in spurious harmonics. Therefore, another novel EMD method was developed for efficient removal of power interference and EMG noise [7]. The reconstruction of the original signal was done by adding up the noise-free IMFs. Safieddine et al. compared EMD with other methods and proved EMD to be best suited for highly contaminated

data [11]. The major drawback of this method was overlapping of modes. To avoid this, ensemble-EMD (EEMD) was proposed in which the average of ensembles of trials was taken as the optimum IMF component [12]. As computational complexity is reduced due to the usage of single channel but it neglects the effect of interchannel dependence. To overcome it, a multivariate empirical mode decomposition (MEMD) was proposed for muscle artifact using a few channels [13].

Recently emphasis has been put on the hybridization of two or more methods such that it utilizes the advantages of each method. One such example is BSS-WT which overcomes the disadvantages of BSS-based analysis as enlisted above. Here, the multichannel signals are converted into ICs or CCs followed by wavelet transform implementation to artifactual component. Thresholding is performed on artifactual coefficients to preserve critical neural information of lower amplitude range [14, 15]. In literature, efficient hybrid methods (compared to BSS-WT) are listed to overcome the disadvantages of BSS technique by combining it with EMD or SVM (artifact specific chosen features are fed into an SVM) [16–18]. Similarly, to overcome the drawbacks of WT, a combination of wavelet and adaptive filter has been listed to remove ocular artifacts [8]. Likewise, Navarro et al. utilized a hybrid of EMD—adaptive filter for ECG artifacts removal from EEG signals [19]. These adaptive filters result in more complicated calculations. That is why a fusion of WT and kalman filter has been proposed for ocular artifact removal by Chen et al. [20]. A combination of EEMD and CCA has been proposed by Sweeney et al. for single-channel motion artifact removal [12]. EEMD-CCA outperforms EEMD-ICA (for single-channel data) in terms of computational time making former one suitable for ambulatory applications. Therefore, CCA appeared to be a suitable method for multichannel signals and EEMD-CCA for single-channel muscle artifact removal [6]. Chavez et al. [21] reported a single channel data-driven surrogates-based artifact removal (SuBAR) approach for removal of OA and MA artifact. It combined wavelet

Table 1 Evaluation parameters for qualitative analysis

Parameters	Formula
Mean square error (MSE)	$\frac{1}{N} \sum_{n=0}^{N-1} [x(n) - \hat{x}(n)]^2$
Peak signal to noise ratio (PSNR)	$10\log_{10} \frac{(\max)^2}{MSE}$
Relative root mean square error (RRMSE)	$\frac{\sqrt{MSE}}{\frac{1}{N} \sum_{n=0}^{N-1} x(n)}$
Correlation coefficient (CC)	$\frac{\text{cov}[x(n), \hat{x}(n)]}{\sigma_{x(n)} \sigma_{\hat{x}(n)}}$
Mean absolute error (MAE)	$\frac{1}{N} \sum_{n=0}^{N-1} x(n) - \hat{x}(n) $
Signal to noise ratio (SNR)	$10\log_{10} \frac{[x(n)]^2}{[x(n) - \hat{x}(n)]^2}$

where; $x(n)$ = clean EEG signal, $\hat{x}(n)$ = estimated (denoised) EEG signal

N = total length of signal, \max = maximum amplitude value of $x(n)$

Cov = covariance, σ = standard deviation

transform with a resampling method (for testing linearity) known as surrogate data approach. Better performance was observed in terms of relative-RMSE compared to combination of CCA with advanced version of the EMD. Table 1 depicts various common evaluation parameters used in literature for qualitative analysis.

4 Conclusion

Artifact identification is a significant part of EEG analysis. It is one of the major problems encountered in bio-signal processing. An extensive survey on the existing artifact removal methods and their limitation has been investigated. As EEG signal varies from person to person, none of the presented methods can be considered as an ideal solution for analysis. Artifact management in the classic EEG recordings is still a dynamic area of investigation. There is a need of a hybrid algorithm that detects all types of artifacts by giving threshold tuning facility to users for various applications. This would significantly improve the signal classification process by reducing the chances of neural data exclusion. Additionally, detail study on commonly encountered artifacts and their characteristics need to be done so as to examine their effect on later stages signal analysis. Finally, in future work focus would be on maintaining a trade-off between performance and computational complexity.

References

1. Jung, C.Y., Saikiran, S.S.: A review on EEG artifacts and its different removal technique. *Asia-Pac. J. Converg. Res. Interchang.* **2**(4), 43–60 (2016)
2. Klados, M. A., Papadelis, C., Braun, C., Bamidis, P.D.: REG-ICA: a hybrid methodology combining blind source separation and regression techniques for the rejection of ocular artifacts. *Biomed. Signal Process. Control.* **6**, 291–300 (2011). Elsevier
3. Zachariah, A., Jai, J., Titus, G.: Automatic EEG artifact removal by independent component analysis using critical EEG rhythms. In: International Conference on Control Communication and Computing (2013). <https://doi.org/10.1109/iccc.2013.6731680>
4. Clercq, W.D., Vergult, A., Vanrumste, B., Paesschen, W.V., Huffel, S.V.: Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram. *IEEE Trans. Biomed. Eng.* **53**(12), 2583–2587 (2006)
5. Yong, X., Ward, R.K., Birch, G.E.: Artifact removal in EEG using morphological component analysis. In: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 345–348 (2009)
6. Khatun, S., Mahajan, R., Morshed, B.I.: Comparative analysis of wavelet based approaches for reliable removal of ocular artifacts from single channel EEG. In: IEEE International Conference on Electro/Information Technology (2015). <https://doi.org/10.1109/eit.2015.7293364>
7. Pachori, R.B.: Discrimination between ictal and seizure-free EEG signals using empirical mode decomposition. *Res. Lett. Signal Process.* **2008**(14) (2008). Hindawi
8. Stearns, S. D.: Adaptive Signal Processing. Prentice Hall, Englewood Cliffs (1985)
9. Jung, T., Makeig, S., Humphries, C., Lee, T., McKeown, M.J., Iragui, V., Sejnowski, T.J.: Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* **3**, 163–178 (2000)

10. Zou, Y., Nathan, V., Jafari, R.: Automatic identification of artifact-related independent components for artifact removal in EEG recordings. *IEEE J. Biomed. Health Inform.* **20**(1), 73–81 (2016)
11. Safieddine, D., Kachenoura, A., Albera, L., Birot, G., Karfoul, A., Pasnicu, A., Biraben, A., Wendling, F., Senhadji, L., Merlet, I.: Removal of muscle artifact from EEG data: comparison between stochastic (ICA and CCA) and deterministic (EMD and wavelet-based) approaches. *Eurasip J. Adv. Signal Process.* **2012**(1) (2012)
12. Sweeney, K.T., McLoone, S.F., Ward, T.E.: The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique. *IEEE Trans. Biomed. Eng.* **60**(1), 97–105 (2013)
13. Chen, X., Xu, T.E., Liu, A., McKeown, M.J., Wang, Z.J.: The use of multivariate EMD and CCA for denoising muscle artifacts from few-channel EEG recordings. *IEEE Trans. Instrum. Meas.* **67**(2), 359–370 (2018)
14. Hamaneh, M.B., Chitravas, N., Kaiboriboon, K., Lhatoo, S.D., Loparo, K.A.: Automated removal of EKG artifact from EEG data using independent component analysis and continuous wavelet transformation. *IEEE Trans. Biomed. Eng.* **61**(6), 1634–1641 (2014)
15. Zhao, C., Qui, T.: An automatic ocular artifacts removal method based on wavelet-enhanced canonical correlation analysis. *IEEE 33rd Annual International Conference of the Engineering in Medicine and Biology Society* (2011). <https://doi.org/10.1109/emb.2011.6091040>
16. Lindsen, J.P., Bhattacharya, J.: Correction of blink artifacts using independent component analysis and empirical mode decomposition. *Psychophysiology* **47**, 955–960 (2010)
17. Chen, X., He, C., Peng, H.: Removal of muscle artifacts from single-channel EEG based on ensemble empirical mode decomposition and multiset canonical correlation analysis. *J. Appl. Math.* **2014** (2014). Hindawi
18. Shoker, L., Sanei, S., Chambers, J.A.: Artifact removal from electroencephalograms using a hybrid BSS-SVM algorithm. *IEEE Signal Process. Lett.* **12**, 721–724 (2005)
19. Navarro, X., Poree, F., Carrault, G.: ECG removal in preterm EEG combining empirical mode decomposition and adaptive filtering. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 2012 pp. 661–664 (2012)
20. Chen, Y., Zhao, Q., Hu, B., Li, J., Jiang, H., Lin, W., Li, Y., Zhou, S., Peng, H.: Method of removing ocular artifacts from eeg using discrete wavelet transform and kalman filtering. In: *IEEE International Conference on Bioinformatics and Biomedicine*, pp. 1485–1492 (2016).
21. Chavez, M., Grosselin, F., Bussalb, A., Fallani, F.D.V., Navarro-Sune, X.: Surrogate-based artifact removal from single-channel EEG. *IEEE Trans. Neural Syst. Rehabil. Eng.* **26**(3), 540–555 (2018)

Hierarchy of Sectors in BSE SENSEX for Optimal Equity Investments Using Fuzzy AHP



Kocherlakota Satya Pritam, Trilok Mathur and Shivi Agarwal

Abstract The stock market is considered as one of the most vital components of a free-market economy. The investments of the stock market are associated with a greater amount of risk. Due to the uncertainty associated with stocks, there is a great need to understand the trend in rising and fall of stocks corresponding to a particular sector or an industry to invest profitably. In view of this, authors have taken up a study to identify the best sector in BSE SENSEX for investments. Fuzzy Analytical Hierarchy Process is used to evaluate and study the dominance of various sectors including Automobile, Finance, Information technology, Oil, Pharmaceuticals, and Power. Four crucial derivatives criteria's Return on equity, Book value per share, Price-earnings ratio, Price to book ratio are considered to study the dominance of each sector. The results of this study help in prioritizing the sectors for future investments.

Keywords Fuzzy analytical hierarchy process · Sectors · Satty scale · Performance index of sectors · BSE SENSEX

1 Introduction

Stock equity is considered as one of the crucial parts of any major economy. It plays a powerful role in growth of economy and hence in HDI (Human development index) which keeps the government, nationalized banks, and industries busy observing the trends of markets closely. Equity market is important from industry and investor's sentiments. The stock market is a collection of companies where capital can be traded by an investor in the form of shares and becomes an owner to a certain portion of

K. S. Pritam · T. Mathur (✉) · S. Agarwal
Department of Mathematics, BITS Pilani, Pilani 333031, Rajasthan, India
e-mail: tmathur@pilani.bits-pilani.ac.in; mathur.trilok76@gmail.com

K. S. Pritam
e-mail: kspritam@gmail.com

S. Agarwal
e-mail: shivi@pilani.bits-pilani.ac.in



Fig. 1 Historical SENSEX from Jan. 1998–Jan. 2018

the company and also assures the warrant of settlement. The primary functioning of the stock market is to collect funds and issue shares to investor and act as a common platform to buyers and sellers. The total market capitalization is above \$69 trillion US dollars in 2015 and out of which the United States possess major share around 34% and with a large gap Japan with about 6% share comes second closely followed by United Kingdom with 5% [1]. There are 16 stock exchanges with a market capitalization of more than 1 trillion US dollars and among them, BSE and NSE are from India.

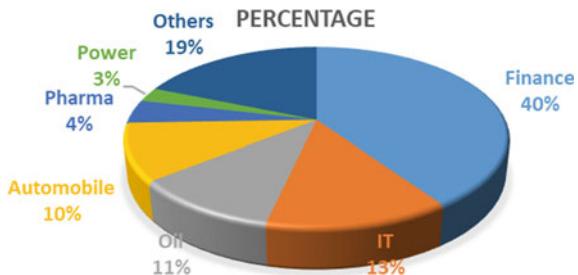
Many stock exchanges are located in India but two are principle due to their significant contribution to Indian economy which are BSE (Bombay stock exchange) and NSE (National stock exchange) located at Mumbai. The present study emphasizes on BSE due to various reasons. The establishment of BSE dates back to 1875 and is the oldest among all the stock exchanges in Asia [2]. By market capitalization, BSE is 11th largest exchange around the globe and with 6 microseconds median trade speed it claims to be the world's fastest stock exchange [1]. There are five indices in BSE which are BSE SENSEX, S&P BSE Small Cap, S&P BSE Mid Cap, S&P BSE Large Cap and BSE 500. Among them the most prominent by capitalization is BSE SENSEX. Hence sectors in BSE SENSEX are given prominence.

Thirty prominent companies are listed in the BSE SENSEX (BSE30) which are listed on the Bombay Stock Exchange. Based on industry representation, the trading volume and liquidity these companies were chosen. BSE SENSEX is capitalization index whose market is of free-float and 100 as its base value. The Index attained its historic high of 36360.22 in January 2018 and marked its least value of 113.28 on December 1979 till January 2018 which can be observed from Fig. 1.

Recently, Indian equity market is performing remarkably even to marginal changes in global markets and its impact is crucial to many major economies. These act as auction between seller and buyer continuously a complying transaction at a location. Protecting the investor, determining realistic price, financing industry, creation of new ventures, attracting foreign investments, and delivering financial needs to the government are some of the objectives of stock exchanges. The growth rate of BSE SENSEX is represented in Table 1.

The volatility of the market is high and investing in the right company is a challenging task. An average investor ends up with losses by trading in the market.

Fig. 2 Sector-wise breakup of SENSEX



Long-term investment of stocks in well-diversified Index funds like BSE SENSEX and NIFTY50 surpassed debt funds since decades [3]. Hence, the investor's portfolio must be diversified and choosing the right company from the sector should be one's priority. Understanding the hierarchy of sectors and investing of sectors in the right proportions is must for a healthy diversified portfolio. There is immense potential growth for the stock market in India and investing in the appropriate sector is a healthy way out. The objective of this study is to tradeoff between profits and reduce the risk. Six sectors are appraised in the present study by their market capitalization and impact to Indian economy which are Automobiles, IT, Oil, Finance, Pharma, and Power. Figure 2 shows the sector-wise breakup of SENSEX.

In the current study, four important financial derivatives Return on equity (ROE), Book value per share (BVPS), Price to earn ratio (PE ratio), and Price to book ratio (PB ratio) are considered which significantly contributes for evaluating the performance of each major sectors. In this study, Fuzzy AHP serves the purpose for prioritization along with equitable portfolio investment.

2 BSE SENSEX Potential

It is evident from the statistics of past years that the number of investments and number of investors in the Indian stock market has increased tremendously. This is due to the high returns on investments. The average annual returns of BSE SENSEX since 2012 is around 13% which is presented in Table 1 and is superior as compared to other debt funds like fixed deposits in nationalized banks etc.

Since the foundation of SENSEX, there is a mammoth variance in the behavior of the stock market with respect to normal distribution. This variation along with other statistical measures are used to scrutinize the behavior of the stock market.

Table 1 Growth rate of BSE SENSEX over 2012–2018

Year	2012–13	2013–14	2014–15	2015–16	2016–17	2017–18
Growth rate	7.58	11.63	39.02	-19.01	21.89	17.46

Table 2 Statistical measures of six sectors

	Mean	S.D.	Var	Min	Max	Range	Skew	Kurt	A-sq
(S1) Finance	885.0	280.9	78942.0	453.8	1581.1	1127.3	0.42	-0.89	32.72
(S2) Auto- mobiles	2038.2	715.5	511914.9	1052.9	3759.0	2706.0	0.45	-0.86	34.01
(S3) IT	1551.5	318.3	101334.1	1156.1	2545.5	1389.4	1.10	0.13	86.85
(S4) Oil	611.7	81.9	6711.1	462.1	908.7	446.6	1.13	1.43	33.04
(S5) Pharma	1268.6	317.0	100461.1	738.5	1999.5	1261.1	0.22	-0.92	17.32
(S6) Power	146.8	23.7	565.1	103.0	201.1	98.1	0.64	-0.68	51.04

The mean, standard deviation, variance, skewness, kurtosis along with 1st quartile, median, and 3rd quartile are well known from literature [4, 5]. These measures unveil the probable risk in the equity market. Further, by using Anderson-Darling normality test, A-squared and p-value are evaluated [6, 7]. All the performance indicators are analyzed and are shown in Table 2.

3 Performance Index Evaluation of Sectors in BSE SENSEX

Evaluating and forecasting equity sectors are determined by the assessment of certain evaluative criteria (synonymously used for financial derivatives). This study considers the following four criteria for analyzing sectors:

Return on equity (ROE-E₁). This is the ratio that measures the return rate interest of investors. It is one of the crucial financial ratios. This criterion can measure the impact of net assets for generating profits and shows the productivity of the company.

$$\text{ROE} = (\text{Net companies equity} - \text{Equity preferred}) / \text{Total companies} \quad (1)$$

Book value per share (BVPS-E₂). This derivative may be employed as a tool to govern sectors equity with respect to the present value of a sector (stock price).

$$\text{BVPS} = \text{Total equity held by the investor} / \text{Total number of shares} \quad (2)$$

Price to earn ratio (PE ratio-E₃). It is a predominant financial derivative to assess a company valuation. This ratio evaluates earnings gained per share. The mathematical equation of PE ratio can be given as follows:

Table 3 Sector-wise values obtained over different criteria

	ROE (%)	BVPS	PE ratio	P/BV
Finance	10.19	10.26	44.77	2.93
Automobiles	20.61	9.64	22.91	4.64
IT	23.22	15.30	16.48	4.04
Oil	10.89	12.65	13.81	1.47
Pharma	14.56	10.25	30.94	3.34
Power	13.54	11.90	13.67	1.86

$$\text{PE ratio} = \text{Worth of a company per share}/\text{Earnings (in total) per share} \quad (3)$$

Generally, investors endeavor to estimate the growth or predict if a company is undervalued or overvalued based on erstwhile trends by using PE ratio [8].

Price to book ratio (PB ratio-E₄). PB ratio is a convenient tool for valuating sectors or companies which obey homogeneous valuations of an asset. Investors consider historical data for predicting a rise in asset price. It can be defined as the ratio between current asset cost in the market and the value of asset (net). This ratio can be explicitly given:

$$\text{PB ratio} = (\text{Current asset cost in the market})/(\text{Assets} - \text{Liabilities in total}) \quad (4)$$

4 Fuzzy Analytical Hierarchy Process (FUZZY AHP)

Even though the statistical measures have significance, there are many disadvantages to their account where the stock market shows positive trends like positive skewness, right kurtosis, etc. But the market crashed [9]. To improve the precision in projection, Fuzzy AHP is used for greater understanding of the market.

For the past several years, multiple criteria decision-making under uncertain environment became a choice in making an appropriate decision. Numerous techniques such as AHP, Electree, extent analysis, DEA have evolved and been used for many applications [10–13]. But many of these methods are proved to end with undesirable results. To overcome these drawbacks and to provide reasonable and logical method Fuzzy AHP is introduced [14–16]. Fuzzy AHP is derived based on AHP in an uncertain environment and decision outcomes. The crisp scale introduced by Saaty is employed onto AHP. Further, Fuzzy AHP redeems crisp Saaty scale to fuzzy scale by using triangular membership function which is utilized in analyzing sectors in the study. The dominance of each sector over the other is measured using fuzzy Saaty scale under multiple evaluative criteria. The schematic representation of the present study is shown in Fig. 3.

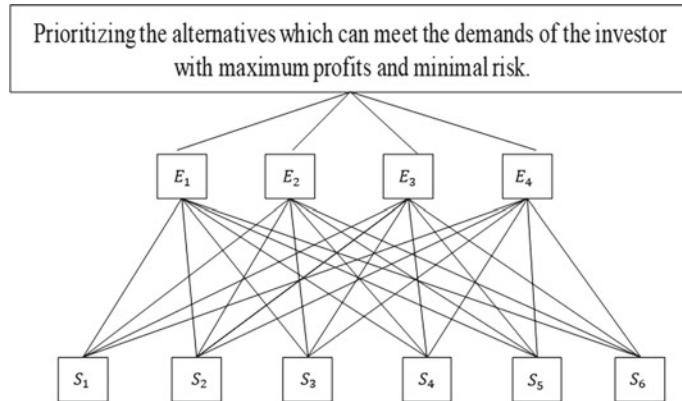


Fig. 3 Pictorial representation of impact of sectors by taking 4 criteria

Four criteria which were considered are discussed in Table 3 with their physical significance. Further, the predominance of each criteria is shown in Table 3.

From the statistics presented in Table 2, it is evident that the attributes are crisp and hence are ineffective for dealing with real-life applications. Since a wide range of criteria are considered in evaluating the sector, weightage for each of the criteria is considered. In the current study, each sector's cumulative score is calculated by using Fuzzy AHP. The advantage is not only restricted to the order of priority but also conveys the optimal investment percentage an individual can invest in a particular sector. Evaluation of the performance index is discussed in the subsequent section.

5 Analysis and Discussion

By using Fuzzy AHP, the outcome of the defuzzified score of each sector is analyzed against the four criterias and then the Performance Index is executed. These five steps are given below.

5.1 Fuzzification of the Crisp Saaty Scale

The range of Saaty scale [17] which is fuzzified vary from 1 to 9 and then fuzzified is shown in Table 4. For fuzzifying the crisp values, triangular membership function is used which represents the generalized fuzzy triangular membership function, which is well known from literature [18]. By analyzing the data from Table 4, the crisp rating is not deviated significantly. Based on this analysis, a span of 1 is preferred in the study for fuzzification (1, 1, 2) along with (8, 8, 9) are considered for the

Table 4 Saaty's crisp and fuzzified scale [17]

Saaty's crisp values (x)	Definition for judgment	Fuzzified Saaty's value
1	Negligible dominance	(1, 1, 1 + η)
3	Dominance is weak	(3 - η, 3, 3 + η)
5	Dominance is strong	(5 - η, 5, 5 + η)
7	Dominance is demonstrative	(7 - η, 7, 7 + η)
9	Dominance is absolute	(9 - η, 9, 9)
2, 4, 6, 8	Intermediate values	(x - 1, x, x + 1)

Table 5 Criteria evaluation (crisp rating)

	ROE (%)	BVPS	PE ratio	PB ratio
ROE (%)	1.00	2.00	0.14	0.17
BVPS	0.50	1.00	0.13	0.14
PE ratio	7.00	8.00	1.00	1.00
PB ratio	6.00	7.00	1.00	1.00

corresponding fuzzification of crisp ratings of border values, respectively. The Saaty scale which is fuzzified related to the prevailing crisp values which can be derived from taking the value of η as 1 where η is the offset distance ranging from 0.5 to 2 [19].

5.2 Analyzing Each Individual Criterion Over the Other

Analyzing the domination of an individual criterion over another is exercised initially to prioritize sectors which are determined by the decision-maker as shown in Table 5. Saaty's crisp scale is utilized to decide the significance of every individual criterion over other. (i.e., range of weights are from 1 to 9) and then assigned the crisp ratings which can be shown from Table 5.

Generalized form of the matrix with respect to criteria evaluation can be given mathematically by (5). The evaluation of criteria is represented generally in the form of matrix shown as shown

$$E = \begin{pmatrix} E_{11} & \dots & E_{1N} \\ \vdots & E_{ij} & \vdots \\ E_{N1} & \dots & E_{NN} \end{pmatrix} \quad (5)$$

where E_{ij} can be interpreted as the dominance of i th evaluative criteria over j th criteria which implies $E_{ij} = 1$ for $i = j$; (where $i = j = 1, 2, \dots, N$) which can be observed

Table 6 Fuzzified weights of criteria

	E ₁			E ₂			E ₃			E ₄		
	a	b	c	a	b	c	a	b	c	a	b	c
E ₁	1.00	1.00	2.00	1.00	2.00	3.00	0.12	0.14	0.17	0.14	0.17	0.20
E ₂	0.33	0.50	1.00	1.00	1.00	2.00	0.11	0.13	0.14	0.12	0.14	0.16
E ₃	6.00	7.00	8.00	7.00	8.00	9.00	1.00	1.00	2.00	1.00	1.00	2.00
E ₄	5.00	6.00	7.00	6.00	7.00	8.00	1.00	1.00	2.00	1.00	1.00	2.00

to be diagonal entries in E matrix. We obtain fuzzified dominance values of E_{ij} by using fuzzified Satty's values shown in Table 6.

Weight of evaluative criteria can be derived by a fuzzy synthetic approach which can be mathematically expressed as

$$w_i = \sum_{j=1}^N s_{ij} \otimes \left[\sum_{k=1}^N \sum_{l=1}^N s_{kl} \right]^{-1}, \quad i = 1, 2, \dots, N \quad (6)$$

where w_i is a fuzzy number which is normalized with unity as its median and s_{ij} are the fuzzified values of E_{ij} obtained from Table 6.

5.3 Analyzing Each Individual Sector Over the Other

Six sectors are considered namely Automobiles, IT, Oil, Finance, Pharma, Power and are evaluated over each other by using four criteria. The mathematical representation can be given as

$$W_k = \begin{pmatrix} b_{11} & \dots & b_{1B} \\ \vdots & b_{ij} & \vdots \\ b_{N1} & \dots & b_{NB} \end{pmatrix} \quad (7)$$

here N denotes the number of criteria and B represents the number of alternatives evaluated over criteria. In Table 7 values are taken into account to correspond all sectors against each of the evaluative criteria which should be normalized initially. Then by using these normalized values, a pairwise comparative matrix is prepared by taking all the sectors and evaluated criteria into consideration.

After the formation of pairwise comparative matrices using the scale mentioned in Table 4, the conversion of crisp ratings to triangular fuzzy number is followed.

Table 7 Sector-wise analysis over criterion 1

	S ₁	S ₂	S ₃	S ₄	S ₅	S ₆
S ₁	1.00	0.13	0.11	1.00	0.25	0.33
S ₂	8.00	1.00	0.50	7.00	5.00	5.00
S ₃	9.00	2.00	1.00	9.00	6.00	7.00
S ₄	1.00	0.14	0.11	1.00	0.33	0.50
S ₅	4.00	0.20	0.17	3.00	1.00	1.00
S ₆	3.00	0.20	0.14	2.00	1.00	1.00

Table 8 Weights of Sectors over E₁

E ₁	a	b	c
S ₁	0.025265	0.03392	0.073478
S ₂	0.211169	0.318829	0.460918
S ₃	0.274205	0.409064	0.54734
S ₄	0.026658	0.037145	0.083426
S ₅	0.069119	0.112697	0.19373
S ₆	0.050038	0.088346	0.164447

Furthermore, by fuzzy division operation, each sector against all the criterions is performed. The weights evaluated against E_1 can be observed from Table 8. The evaluated weights for the other three criteria can be evaluated similarly, i.e., E_2 , E_3 , and E_4 .

5.4 Computing a Performance Matrix and Determination of the Score Which Is Fuzzified

Fuzzy interval arithmetic is applied for the sector-wise performance of each of the 6 sectors over throughout the 4 criterions which can be illustrated through the performance matrix [20] given by (8).

$$z = \begin{bmatrix} y_{11} \otimes w_1 & \dots & y_{1k} \otimes w_k \\ \vdots & \ddots & \vdots \\ y_{N1} \otimes w_1 & \dots & y_{Nk} \otimes w_k \end{bmatrix} \quad (8)$$

where $\{y_{11}, \dots, y_{Nk}\}$ are set of priority vectors and $\{w_1, \dots, w_k\}$ represents weights of each criteria. Evaluation of performance matrix z corresponding to each of the sectors over E_1 , E_2 , E_3 , and E_4 are calculated. Table 9 indicates the performance matrix of sector S_1 against the criteria. The performance matrix of the remaining sectors (S_i 's) can be evaluated similarly.

Table 9 Performance matrix for sector S_1 (finance)

P_1	a	b	c
E_1	0.025265	0.03392	0.073478
E_2	0.035583	0.047475	0.106652
E_3	0.336216	0.494761	0.634036
E_4	0.085424	0.139965	0.234956

Table 10 Overall sector performance evaluated against 4 criteria

Sectors	Fuzzy weight of sectors (overall)		
S_1	0.424822	0.199718	0.095585
S_2	0.450793	0.197071	0.08407
S_3	0.574448	0.288236	0.133841
S_4	0.190956	0.078635	0.03565
S_5	0.375517	0.160139	0.070376
S_6	0.196684	0.0762	0.031891

The summation of all the assessments maximum, minimum, median, mean, decision-maker and mixed operators are utilized. The current study adopts additive synthesis to analyze the performance for each Sector which means to reduce risks or negative impact and to maximize returns. The sector-wise performance is given as:

$$F_i = \sum_{j=1}^c y_{ij} \otimes w_j \quad (9)$$

In (9) F_i represents fuzzy score of each Sector; y_{ij} represents priority vectors and w_j represents evaluated criteria's fuzzy weight (Table 10).

5.5 Defuzzifying Cumulative Score of Each Sector and Computing the Values

The aggregated score for all the sectors is considered for the sector-wise prioritization. As discussed about calculation of score in Sect. 5.4. Each value of the sectors is in triangular fuzzy number. Hence for converting this fuzzy number to a crisp value, defuzzification of each sector is done. There are many methods to determine the defuzzified score by the level of optimism, in this study the method of total integral value is considered for a wide range of benefits [21]. The estimation of a fuzzy number $[a, b, c]$ by using the method of total integral value can be mathematically given as:

Table 11 Scores of sectors which are defuzzified with respect to specified λ

	$\lambda = 0$	$\lambda = 0.5$	$\lambda = 1$	Ranking
S ₁	0.239305	0.174449	0.478611	3
S ₂	0.248428	0.176461	0.496855	2
S ₃	0.336948	0.247883	0.673896	1
S ₄	0.106747	0.07501	0.213495	5
S ₅	0.205555	0.145593	0.411109	4
S ₆	0.107185	0.07398	0.214371	6

$$I_T^A = 0.5[\lambda c + (1 - \lambda)a + b], \lambda \in [0, 1] \quad (10)$$

where, λ is the optimism index, which indicates the decision-maker's level of risk. The value λ is directly proportional to optimism degree. Generally, $\lambda = 0, 0.5, 1$ corresponds to pessimist, moderate, and optimist, respectively. Then the total integral values for each sectors are calculated by (10) for getting defuzzified score corresponding to λ where $\lambda = 0, 0.5, 1$ which can be shown as Table 11.

6 Conclusions

BSE SENSEX is significantly growing year by year and is expected to grow at a much faster pace due to the exceptional GDP growth rate of India and hence its suitable time for investors have a well-diversified sectored portfolio. Prioritizing of sectors in the right proportions is a must, for a healthy portfolio. A vigilant assessment of parameters is made by choosing fuzzy AHP. This study also provides the significance of important parameters that can give more returns and an opportunity to have a healthy portfolio by analyzing the performance of the major six sectors during the past five years. The findings of the study infer that all sectors have positive skewness with IT sector possessing a high value of skewness along with low positive kurtosis and oil sector with high skewness and high kurtosis. Besides these, the rest of the four sectors have negative kurtosis with pharma sector attaining the least. Further, it is observed that all the sectors do not follow a particular trend and are highly volatile with IT sector outpaces others in volatility. IT sector is found to be dominant than other sectors by a huge margin followed by automobile and financial sector. The remaining sectors are lagging far behind in performance. Furthermore, the hierarchy drawn, and the relative dominance help for optimal investment can aid proportional investments by integrating different multi-criteria techniques by considering the perception of the investor in future.

Acknowledgements P. Satya would like to thank DST-FIST vide SR/FST/MSI-090/2013(C) for providing infrastructure support and UGC for providing financial support as JRF (Ref. No: 22/12/2013(ii)EU-V with Sr. No. 2121341012).

References

1. Menon, S.: A comparative study of the Indian stock market with TWO international stock markets between 2012 and 17. *5*(3), 14 (2018)
2. Poshakwale, S.: Evidence on weak form efficiency and day of the week effect in the Indian stock market, *12* (1996)
3. Modigliani, F., Miller, M.H.: The cost of capital, corporation finance and the theory of investment. *Am. Econ. Rev.* **48**(3), 261–297 (1958)
4. Cleary, R.J.: Applied data mining: statistical methods for business and industry. *J. Am. Stat. Assoc.* **101**(475), 1317–1318 (2006)
5. McEvoy, D.M.: *A Guide to Business Statistics*. Wiley, Hoboken, NJ, USA (2018)
6. Hadi, A.R.A., et al.: The effect of oil price fluctuations on the malaysian and indonesian stock markets, *23* (2009)
7. Pettitt, A.N.: Testing the normality of several independent samples using the Anderson-Darling statistic. *Appl. Stat.* **26**(2), 156 (1977)
8. Peter, D.: Easton: PE ratios, PEG ratios, and estimating the implied expected rate of return on equity capital. *Account. Rev.* **79**(1), 73–95 (2004)
9. Campbell, J., Hentschel, L.: No News is Good News: an Asymmetric Model of Changing Volatility in Stock Returns. National Bureau of Economic Research, Cambridge, MA (1991)
10. Agarwal, S.: Scale efficiency with fuzzy data. *Int. J. Bus. Syst. Res.* **11**(1–2), 152–162 (2017)
11. Erensal, Y.C., et al.: Determining key capabilities in technology management using fuzzy analytic hierarchy process: a case study of Turkey. *Inf. Sci.* **176**(18), 2755–2770 (2006)
12. Jha, S.K., Puppala, H.: Prospects of renewable energy sources in India: prioritization of alternative sources in terms of energy index. *Energy* **127**, 116–127 (2017)
13. Kahraman, C. (ed.): *Fuzzy Multi-criteria Decision Making: theory and Applications with Recent Developments*. Springer, New York, NY (2008)
14. Chatzimouratidis, A.I., Pilavachi, P.A.: Multicriteria evaluation of power plants impact on the living standard using the analytic hierarchy process. *Energy Policy* **36**(3), 1074–1089 (2008)
15. Dağdeviren, M., Yüksel, İ.: Developing a fuzzy analytic hierarchy process (AHP) model for behavior-based safety management. *Inf. Sci.* **178**(6), 1717–1733 (2008)
16. Kalban, M.M.: Decision support for energy conservation promotion: an analytic hierarchy process approach. *Fuel Energy Abstr.* **45**(5), 365 (2004)
17. Saaty, T.L.: Decision making with the analytic hierarchy process. *Int. J. Serv. Sci.* **1**(1), 16 (2008)
18. Ertuğrul, İ., Karakaşoğlu, N.: Performance evaluation of Turkish cement firms with fuzzy analytic hierarchy process and TOPSIS methods. *Expert Syst. Appl.* **36**(1), 702–715 (2009)
19. Jain, V., et al.: Supplier selection using fuzzy AHP and TOPSIS: a case study in the Indian automotive industry. *Neural Comput. Appl.* **29**(7), 555–564 (2018)
20. Seçme, N.Y., et al.: Fuzzy performance evaluation in turkish banking sector using analytic hierarchy process and TOPSIS. *Expert Syst. Appl.* **36**(9), 11699–11709 (2009)
21. Liou, T.-S., Wang, M.-J.J.: Ranking fuzzy numbers with integral value. *Fuzzy Sets Syst.* **50**(3), 247–255 (1992)

A Novel Inset Fed Patch Antenna with Defecting Ground Plane



Pratik Ghosh, Kousik Roy, Atanu Nag, Rashmi Priya and Naimul Hasan

Abstract This paper illustrates the constructional plan of an inset feed mechanism microstrip patch antenna. The essential advantage of employing inset cut feed is not only increasing the antenna bandwidth but its return loss is also improved suitably. Our antenna has been simulated using hexagonal patch on both defected and non-defected ground plane. It has been realized that operation of the antenna, acquires development according to the size of the patch, over usual design.

Keywords Inset feed · Return loss · Bandwidth · DGS

1 Introduction

Microstrip Patch Antenna (MPA) is one of the utmost accepted one due to its various advantages like miniature size, low profile, less weight, simple structure, low cost, simple fabrication, and installation techniques, etc. [1–3] with the noteworthy shortcoming of narrow bandwidth. Out of the various techniques by which antenna bandwidth may be improved, the implemented one is the defected ground structure

P. Ghosh

Murshidabad College of Engineering and Technology, 742102 Murshidabad, WB, India
e-mail: prakanika2014@gmail.com

K. Roy (✉) · R. Priya

Asansol Engineering College, 713305 Asansol, WB, India
e-mail: kousikroy002@gmail.com

R. Priya

e-mail: priyarashmi1992@gmail.com

A. Nag

Modern Institute of Engineering & Technology, 712123 Bandel, Hooghly, WB, India
e-mail: tnnag79@gmail.com

N. Hasan

Sanaka Educational Trust's Group of Institution, 71321 Durgapur, WB, India
e-mail: naimul_hasan@yahoo.in

(DGS) [4, 5]. The limitation produced in a ground plane is a typical characteristic technique for minimizing antenna size. Antenna designed with DGS reduces the antenna size and improves the gain or bandwidth of the antenna. The remaining part of the paper is organized as follows: Sect. 2 describes the design geometry of the proposed antenna, Sects. 3 and 4 describe respectively the simulation results of proposed MPA without and with DGS, Sect. 5 gives their comparative analysis and Sect. 6 concludes the paper.

2 Design Geometry

Figure 1a, b depicts the geometry of proposed MPA with and without DGS. As depicted in figure, inset feed is utilized to feed the MPA of definite feed line width so that the antenna impedance equals with port impedance (50Ω) [6, 7] by maintaining the width and spacing of feed line. The fundamental benefit of utilizing inset cut feed is not only increasing the antenna bandwidth but return loss is also improved satisfactorily. The antenna is planned on 1.57 mm thick substrate with dielectric constant 4.7.

3 Simulation Results of Proposed MPA Without DGS

For simulation of the antenna we have used ANSYS HFSS 2017 and the operation has been studied in terms of return loss, gain characteristics, VSWR, input impedance, current distribution, and 3D polar plot [8–10]. Without DGS we have got the bandwidth and VSWR as 3.07 GHz and 1.1294 respectively.

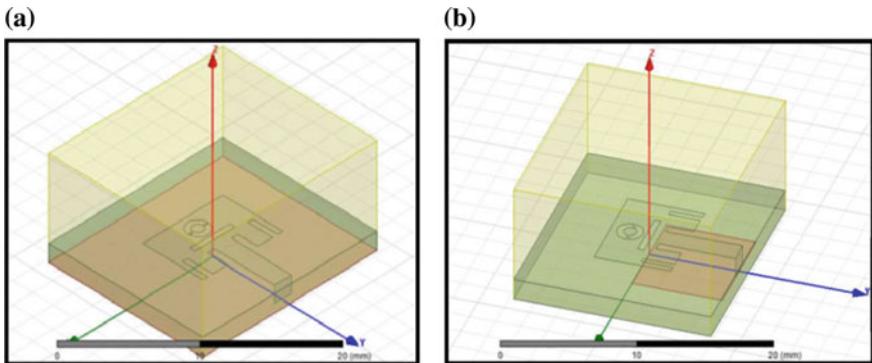


Fig. 1 3D geometry of the proposed MPA (a) without DGS and (b) with DGS

3.1 Return Loss and Bandwidth

Figure 2 shows the simulated return loss and bandwidth of inset fed slot loaded MPA which has a resonance frequency of 10 GHz. It has been examined that the return loss of proposed design is -24.324 dB.

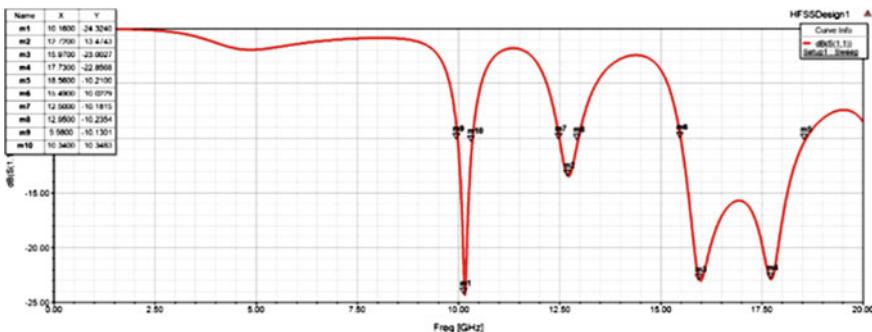


Fig. 2 Return loss and bandwidth of slot loaded MPA without DGS

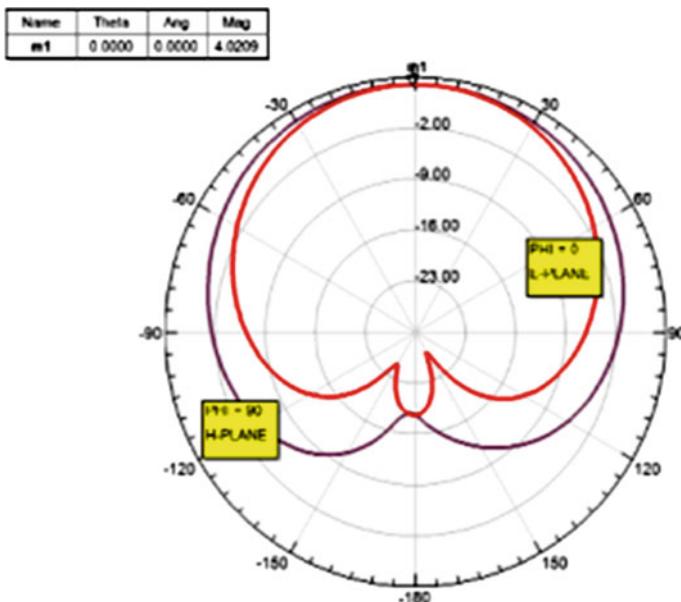


Fig. 3 Gain pattern for inset fed slot loaded MPA without DGS

3.2 Gain, Input Impedance, and VSWR

Figure 3 shows the 3D radiation pattern representing the gain of MPA. The examined gain at resonant frequency is 4.0209 dB. Figure 4 depicts the input impedance of the proposed MPA without DGS. At resonant frequency the analyzed input impedance of the proposed antenna is 46.22Ω . Figure 5 shows the VSWR plot for proposed inset fed slot loaded MPA without DGS.

Name	Freq	Ang	Mag	R.X
m1	10.1600	-127.5556	0.3628	$0.5244 + 0.0854i$

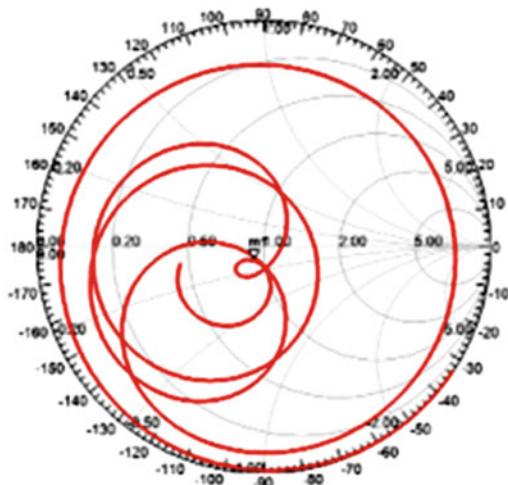


Fig. 4 Input impedance for inset fed slot loaded MPA without DGS

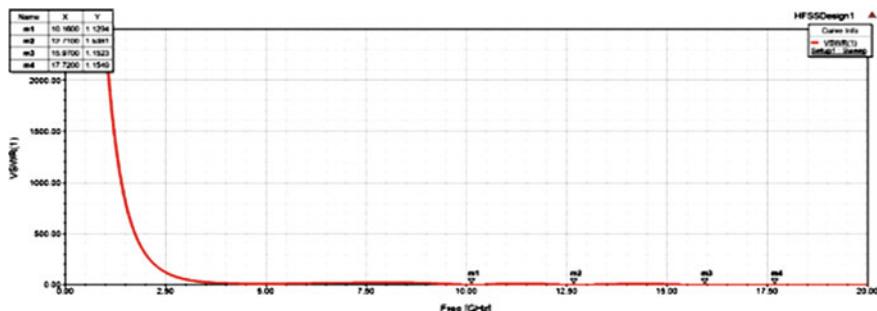


Fig. 5 VSWR for inset fed slot loaded MPA without DGS

3.3 Surface Current Distribution and 3D Polar Plot

Figure 6 depicts the current distribution on the patch of proposed MPA without DGS. Figure 7 depicts the 3D polar plot radiation pattern of proposed MPA without DGS.

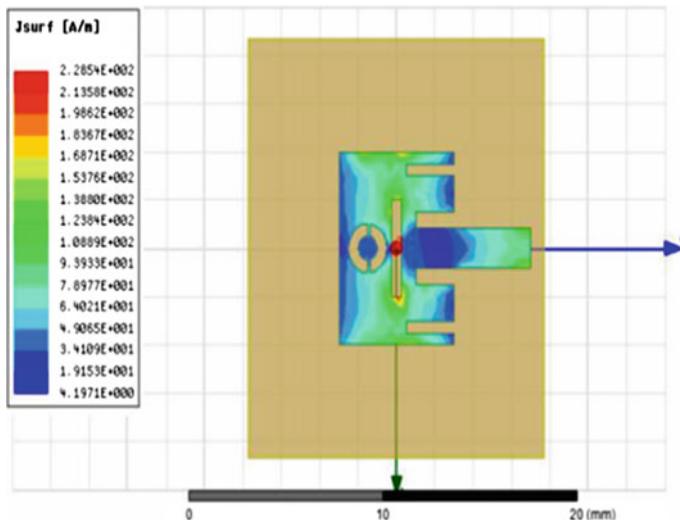


Fig. 6 Surface current distribution for inset fed slot loaded MPA without DGS

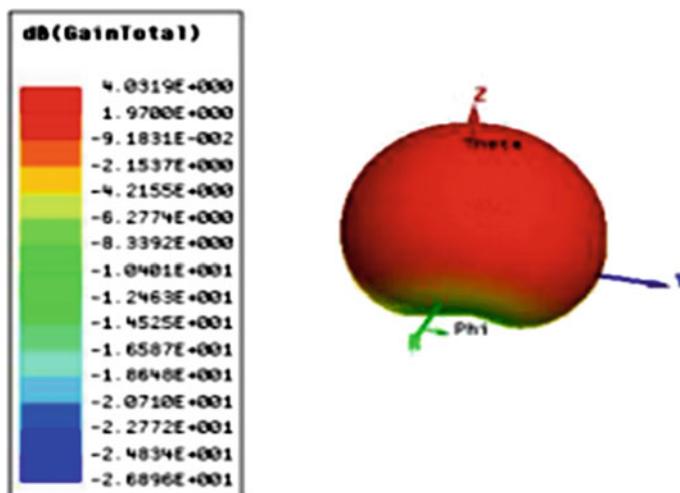


Fig. 7 3D polar plot for inset fed slot loaded MPA without DGS

4 Simulation Results of Proposed MPA Having Hexagonal Slot Cut with DGS

Simulation of the suggested slot loaded inset fed MPA with DGS along with hexagonal slot cut has been obtained using ANSYS HFSS 2017 and the operation of the antenna has been studied typically by analyzing its return loss, gain, VSWR, input impedance, current distribution, and 3D polar characteristics [8–10].

4.1 Return Loss

Figure 8 depicts the simulated return loss as well as bandwidth of inset fed slot loaded MPA with DGS.

4.2 Gain, Input Impedance, and VSWR

Figure 9 shows the 3D radiation pattern representing the gain characteristics of proposed slot loaded MPA having defecting ground plane hexagonal slot cut. Observed gain at resonant frequency is 3.9122 dB. Figure 10 depicts the input impedance of the proposed MPA having hexagonal slot cut with DGS.

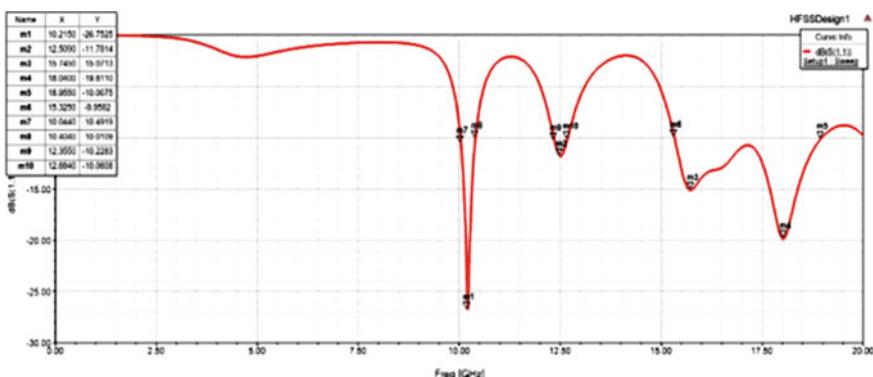


Fig. 8 Return loss and bandwidth of slot loaded MPA having DGS with hexagonal slot cut

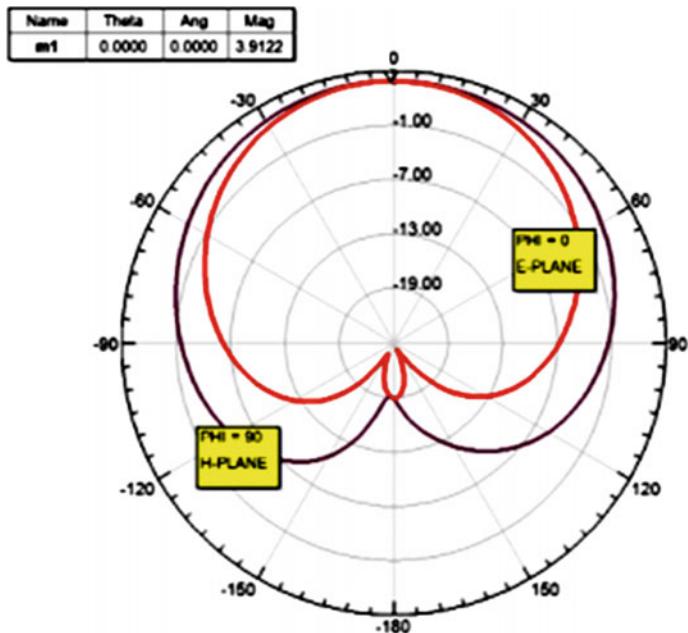


Fig. 9 Gain characteristics of slot loaded MPA having DGS with hexagonal slot cut

Name	Freq	Ang	Mag	RX
m1	10.2080	-139.2107	0.0478	0.9264 - 0.0581i

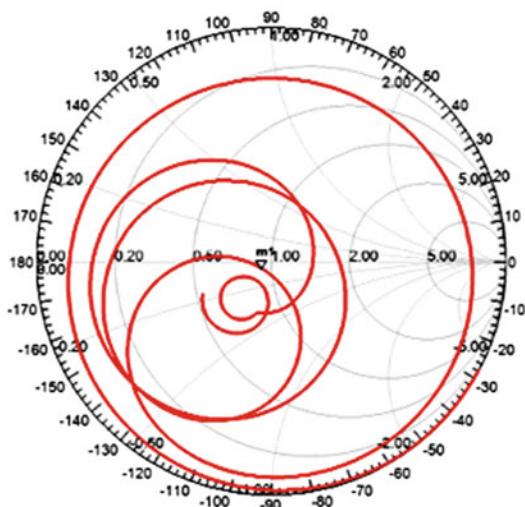


Fig. 10 Input impedance of hexagonal slot cut MPA having DGS

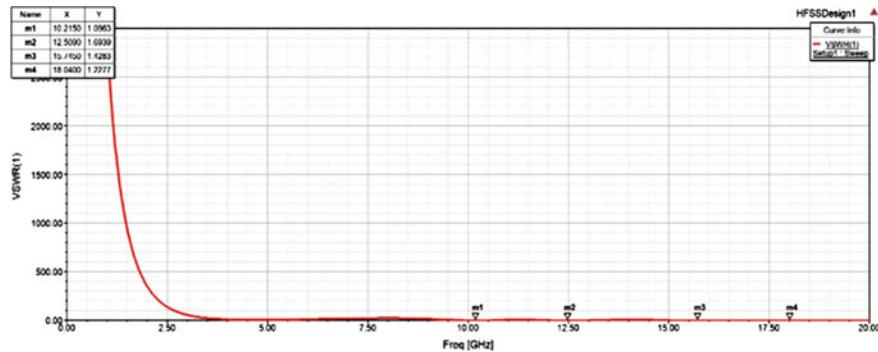


Fig. 11 VSWR result of slot loaded MPA having hexagonal slot cut with DGS

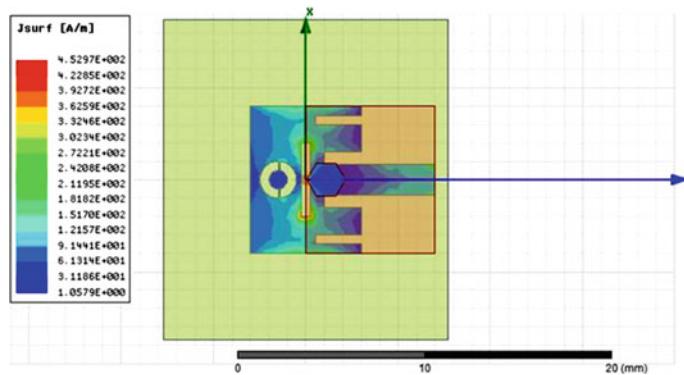


Fig. 12 Current distribution of hexagonal slot cut MPA having DGS

Figure 11 depicts the VSWR plot for proposed slot loaded MPA having hexagonal slot cut with DGS.

4.3 Surface Current Distribution and 3D Polar Plot

Figure 12 depicts the current distribution on the patch of proposed MPA with DGS. Figure 13 depicts the 3D polar plot radiation pattern of proposed MPA with DGS.

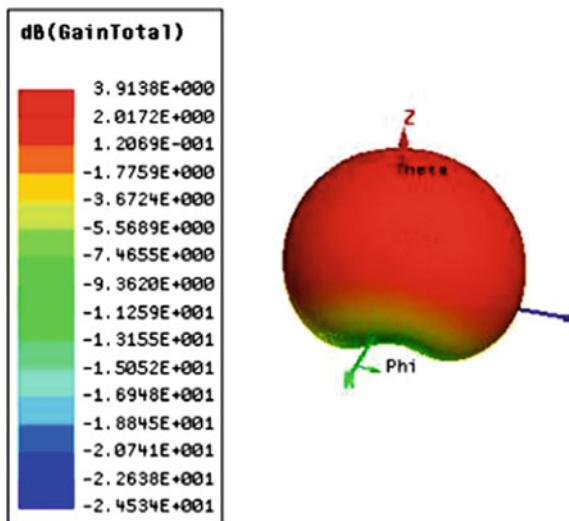


Fig. 13 3D polar plot of gain of hexagonal slot cut MPA having DGS

Table 1 Comparative study of various proposed design of MPA

Proposed MPA	Return loss (dB)	Bandwidth (GHz)	VSWR	Gain (dB)	Input impedance
Without DGS	-24.3240	3.07	1.1294	4.0209	46.22
With DGS	-25.1000	3.6	1.1177	3.9249	47.415
With DGS having hexagonal slot	-26.7525	3.63	1.0963	3.9122	46.42

5 Comparison of Results Between DGS and Non-DGS Proposed MPA

Table 1 illustrates the comparative study of various proposed design of MPA in terms of the analyzed return loss, bandwidth, VSWR, gain, and input impedance.

6 Conclusion

The designed MPA with loaded inset fed with and without DGS have been simulated in ANSYS HFSS 2017 in terms of gain, input impedance, return loss, bandwidth, VSWR, and surface current. It is perceived that bandwidth and return loss has been improved by utilizing DGS technique. Simulated conclusion recommends that the

reliance of the input impedance on feed position is unlike for the defected and non-defected ground plane in company with diverse slot cut. By concerning defective ground plane on the proposed inset fed hexagonal slot cut MPA the bandwidth, return loss, and input impedance get improved adequately.

References

1. Kraus, J.D.: *Antennas for All Applications*, 5th edn. McGraw Hill, Inc., New York
2. Balanis, C.A.: *Antenna Theory: Analysis and Design*, 4th edn. Wiley, New York
3. Tan, B.K., Withington, S., Yassin, G.: A compact microstrip-fed planar dual-dipole antenna for broadband applications. *IEEE Antennas Wirel. Propag. Lett.* **15**, 593–596 (2016)
4. Milligan, A.T.: *Modern Antenna Design*, 2nd edn. Wiley, New York
5. Partha, D.K., Moyra, T., Bhowmik, P.: Return loss and bandwidth enhancement of microstrip antenna using defected ground structure (DGS). In: 2nd International Conference on Signal Processing and Integrated Networks (2015). <https://doi.org/10.1109/spin.2015.7095318>
6. Henry, A.D., Jung, Y.B.: broadband proximity-coupled microstrip planar antenna array for 5G cellular applications. *IEEE Antennas Wirel. Propag. Lett.* **17**(7), 1286–1290 (2018)
7. Rajgopal, K.S., Sharma, K.S.: Investigations on ultrawideband pentagon shape microstrip slot antenna for wireless communications. *IEEE Trans. Antennas Propag.* **57**(5) (2009)
8. Roy, K., Nag, A., Chaudhuri, D., Bose, S.: A novel dual band antenna for C and X band satellite communication. In: International Conference on Electrical, Electronics, Signals, Communication and Optimization, pp. 953–955. IEEE (2015). ISBN: 978-1-4799-7676-8, 978-1-4799-7678-2/15/\$31.00 ©2015
9. Roy, K., Chaudhuri, D., Bose, S., Nag, A.: A novel dual band antenna for radar application. In: Proceedings of 3rd International Conference on Advanced Computing Networking and Informatics, vol. 2, 44, pp. 643–650, Springer, India (2015). ISBN: 978-81-322-2528-7. <https://doi.org/10.1007/978-81-322-2529-4>
10. Nag, A., Roy, K., Chaudhuri, D.: A crown-shaped microstrip patch antenna for wireless communication systems. In: Kalam, A., Das, S., Sharma, K., (eds) *Advances in Electronics, Communication and Computing, Lecture Notes in Electrical Engineering*, vol. 443, pp. 399–405 (2018)

Bionics Based Camouflage Clothing for Soldiers



Shweta Gupta, S. A. Hariprasad and Adesh Kumar

Abstract This research paper emphasizes on the bionics based colour changing clothing for camouflage of soldiers which changes colour based on the concept used in chameleon and octopus. It is made up of smart thread which is coated with photochromic paint and with the change in light it changes colour as in case of chameleon. And the other type of clothing is based on Octopus skin which changes its colour based on temperature variation. This clothing can be used for military purposes or for changing the colour of clothing of groom and bride in wedding ceremony. The application can be further enhanced to changing colour of the day today clothing depending on the mood of person wearing the clothing. As the person enters the office his clothing colour could be changed to some formal colour and as he enters out of office in evening to the lighter different colour. Thus, instead of having fifty different T-Shirts, he can have only one T-shirt, and further, this concept can be further extended to fashion industry.

Keywords Chameleon · Octopus · Bionics · Photochromic paint · Camouflage · Smart thread

1 Introduction

Bionics (also referred as biomimetics, biomimicry, biognosis or bionical creativity engineering) involves applying the possible applications of biological methods and systems found in nature to the studying and designing of engineering systems and modern technology. Bionics examples include imitating the thick skin of dolphins

S. Gupta (✉) · S. A. Hariprasad
Electronics and Communication Engineering Department, Jain University, Bangalore, India
e-mail: shwetagupta832000@gmail.com

A. Kumar
Department of Electrical and Electronics Engineering, School of Engineering,
University of Petroleum and Energy Studies, Dehradun, India
e-mail: adeshmanav@gmail.com

Fig. 1 Chameleon
(*Furcifer Pardalis*) in angry
excited state [5]



for making hulls of boats; and echolocation of bats for sonar, radar and medical ultrasound imaging. A remarkable feat in the field of computer science involves bionics based artificial neurons, artificial neural networks and intelligence [1].

The strange ability of chameleon to change colour has been a mystery for a long time but now the mystery has been revealed. The scientific reason behind chameleon changing colours involves changing the width of the layer of special cells. Besides, comparing with other animals which change colour like squid and octopus, the phenomenon behind chameleons changing colour is how light reflects off from their skin depending on the structural changes [2].

After studying adult male five in number, adult female four in number and adult panther chameleons four in number (*Furcifer Pardalis*), a species found in Madagascar, it was concluded from the findings that there have superimposed two thick layers of cells of iridophore. The concluded iridescent cells contain a pigment which eventually results in reflection of light. Nanocrystals are contained in iridophore cells [3] which involve the reflection of light of longer wavelength such as yellow, orange, red and yellow mixed with blue helps the chameleon to hide in plants and trees.

Chameleon changes colour to camouflage itself and the same concept can be used to develop the clothing for soldiers using bionics to camouflage themselves during war. The more elaboration upon on how chameleon changes colour is they have special type of cells called Chromatophores. The manipulation of these cells results in the change of colour. Chameleons have two superimposed layers within their skin, of which upper layer consists of different sized nanocrystals. A change in the size and shape of these nanocrystals [4], in the relaxed state or excited state of the chameleon, results in the change of colour. The relaxed mode of skin of chameleon makes the iridophore cells in the nanocrystals to come close to each other, and thus result in reflection of short wavelength such as blue. While, in the excited state of chameleon skin, the distance between the nanocrystals in the iridophore cells increases and each of these iridophore cells (which contains these nanocrystals) results in reflection of longer wavelengths, such as yellow, orange and red. The Chameleon (*Furcifer*



Fig. 2 Chameleon (*Furcifer Pardalis*) changes colour to camouflage [5]

Pardalis) in angry excited state is shown in Fig. 1 and behaviour of Chameleon (*Furcifer Pardalis*) changes colour to camouflage is depicted in Fig. 2.

These cells act as prism and reflect light of different wavelengths.

2 Inspiration from Nature

Till date, we are making displays like LED, LCD and CRT that are not so flexible, on the contrary are brittle, rigid and bulky and can break down. Even displays in mobile phones require light source, filters and glass plates [6]. As compared to the animals like octopuses, chameleons and squids have skin like displays, which are flexible, thin and can change colours by birth [7].

Thus, an inspiration from Nature to make ultra-thin, colour changing clothing not only for camouflage for soldiers but also to have colour changing one T-Shirt in wardrobe instead of having 50 shirts. Even, the groom could change the colour of his clothing as he steps towards his bride in wedding with colours of wedding hall.

3 Clothing Material for Camouflage

A clothing that could blend in the background. A wonderful feat for students who have not done their homework and when they are called by their teacher, they could blend into desk or chair without being noticed. Thus, a clothing material that changes its appearance based on the surroundings. Thus, a material with specialized property of changing colour, with special type of thread containing crystal, that can change its shape and colour when different wavelengths of light are shone on it, which could be used for active camouflage [8].

The basic technology involved smart thread coated with photochromic paint which changes colour when exposed to light. A pixel of flexible sheet that is light sensitive and changes its pattern to match. If the material expands, the pixel changes colour

and if it contracts, it comes back to normal. These materials function by changing colours, when light falls on it the photochromic crystals come together as shown in Fig. 3 and coming back to original color, when the light is removed from it and becomes ‘normal’. Material is sensitive to light, which being its key characteristic of changing colour, flexible, waterproof and scratch-resistant [9].

These threads are made of non-phthalate polyvinyl chloride, thermoplastic polyurethane, polypropylene, polyethylene, thermoplastic olefin (TPO) and ther-

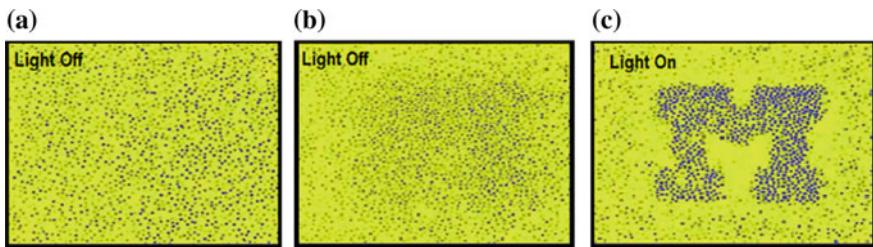


Fig. 3 Smart material changes colour when light falls on it [10]

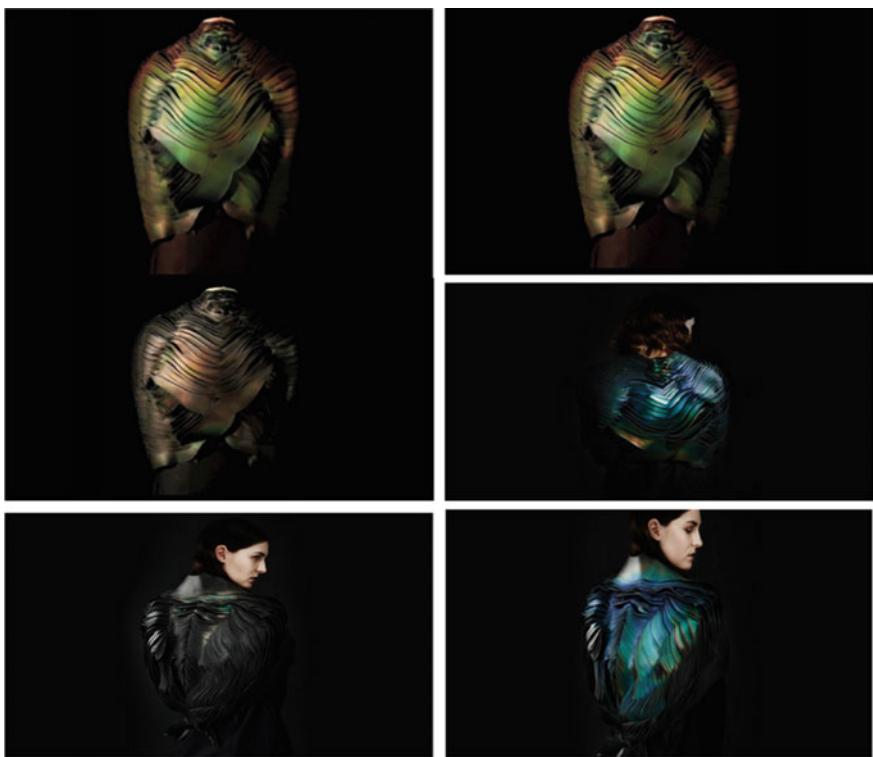


Fig. 4 Colour changing clothing with its application in fashion [14]

moplastic elastomers [11]. Photochromic implies that the fabric changes its colour when exposed to sunlight within the range of 300–360 nm. Photochromic pigments are coated around the threads and can easily be weaved and have aesthetic effects and are power efficient. The photochromic pigments slowly change colours and when woven into fabrics result in calming ‘animations’ which results in moving across the fabric [12].

The clothing made from these materials can easily be washed without the fear of being scratched off. These fabrics can be cut to produce different types of clothing, sewn, welded in clothing and printed. They are lightweight and portray to be antennas that can be washed [13]. Figure 4 presents the colour changing clothing with its application in fashion.

4 Material Based on Octopus Skin

The material consists of squared grid of 16 by 16 grid which is layered based on octopus skin which has three layers. Topmost one contains cells called chromatophores, which are sacs of coloured pigments, controlled by ring of muscles. On the expansion of sac, it produces pixel of colour and if it contracts, pixel hides. Cells have the main responsibility of hues like red, orange, yellow and black. Middle layer contains iridophores, cells that reflect light and is responsible for cooler colours like blues and greens. The bottom layer consists of leucophores which are passive cells and involves diffusion of light in all directions. Octopus skin also contains opsins which are light sensitive molecules and are majorly found in retina.

So, based on that we came up with material in which,

1. The topmost layer contains heat sensitive dye which involves changing colour from black to room temperature to colourless at 47 degree Celsius and back again like the chromatophores in octopus skin.
2. The next layer involves a thin piece of silver which gives rise to bright white background like leucophores.
3. Below it is the diode which helps in heating the overdyeing dye and control the colour. This is analogous to muscles that control chromatophores.
4. Finally, there is a layer which consists of light detector in the corner, a bit like skin of cephalopod’s containing opsins.
5. The entire structure rests on a base that is flexible and can blend and flex without breaking.

The conclusion of how the material is exactly working is the light detector sense any incoming light and the diodes in illuminated panel to heat up and thus the colour of the overdyeing dye changes. The pixels reflect light from the silver layer giving the white look and pixels are shown in Fig. 5 as clothing material changing colour. In the same way, Fig. 6 presents the example of Bed as changing colour.

The technology is further enhanced to metal micro-wires embedded in each individual thread and eventually controlled by smartphone. All this happens with

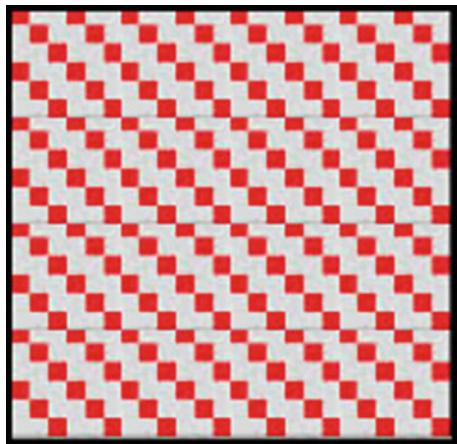


Fig. 5 Clothing material changing colour



Fig. 6 Bed as changing colour



Fig. 7 Smartphone controlled fabric



Fig. 8 Micro thin metal wires with the coating

smartphone-controlled colour changing fabric with the help of an app, it changes with the concept of optics and photonics. Figure 7 shows how smartphone with app can be applicable for changing the colour. Figure 8 shows the different Micro thin metal wires with the coating.

Upon activation, slight electric current flow through the wires, leading to the increase in temperature and resulting in special pigments in thread to change colour. Similar application could be in wedding parties in which colour of the clothing of groom could be changed and can further be extended to button controlled fabric. The spinning metals thread embedded in fabric is shown in Fig. 9. The fabric can be controlled with the help of push button as shown in Fig. 10. In this way different colours can be obtained over the fabric.

5 Applications

Various applications of colour changing clothing based on chameleon skin and octopus skin have been discussed and it can be used in soldiers clothing for camouflage in wars, etc. Further, they can be used in normal person in day today life, like instead of having 50 different coloured T-shirts in the wardrobe, we can come up with single shirt that can change colour depending on your mood and as you move into your

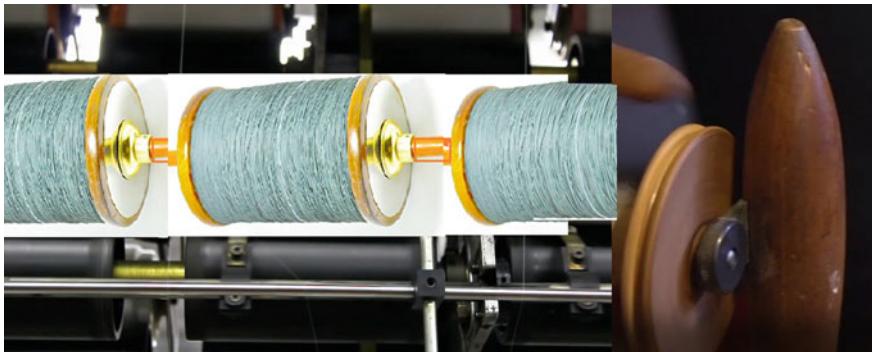


Fig. 9 Spinning metal thread to be embedded in fabric



Fig. 10 Button controlled fabric

office, it takes a formal colour, as you move out from the office in evening to a restaurant, it takes the different colour. But the fabric based on the chameleon skin is much better because it is light sensitive and changes colour because of light exposure while based on octopus skin it changes colour using electric current to change temperature.

6 Conclusion

These kind of fabrics which changes colour on exposure to light are not only lightweight or the colour change can be controlled by button are very useful not only for camouflage but also for fashion industry. With its application in fashion industry, the application can also be extended to soldiers in the heights of mountains and protecting the lives of civilians. Though besides, photochromic fabric, other

ways like thermochromic fabric is there but photochromic fabric is the most effective way of colour changing fabric which ensembles to be the fabric of next era and has innumerable applications.

Disclaimer Authors would like to state that Chameleons (Furciper Pardalis) are not directly used in this study.

References

1. Gérardin, L.: *Bionics*. McGraw-Hill, New York (1968)
2. The chameleon—how it changes color. *Sci. Am.* **7**(157), 2502–2503 (1879)
3. Goda, M.: Rapid integumental color changes due to novel iridophores in the chameleon sand tilefish Hoplolatilus chlupatyi. *Pigment Cell Melanoma Res.* **30**(3), 368–371 (2017)
4. Service, R.: The secret to chameleon color change: tiny crystals. *Science* (2015)
5. <https://www.youtube.com/watch?v=SQggDnScsvI>
6. Gao, X., Guo, Z.: Biomimetic superhydrophobic surfaces with transition metals and their oxides: a review. *J. Bionic Eng.* **14**(3), 401–439 (2017)
7. Sakai, M., Seki, T., Takeoka, Y.: Bioinspired color materials combining structural, dye, and background colors. *Small*, 1800817 (2018). <https://doi.org/10.1002/smll.201800817>
8. Franklin, D., Chen, Y., Vazquez-Guardado, A., Modak, S., Boroumand, J., Daming, X., Shin-Tson, W., Chanda, D.: Polarization-independent actively tunable colour generation on imprinted plasmonic surfaces. *Nature Commun.* **6**, 7337 (2015). <https://doi.org/10.1038/ncomms8337>
9. Structures, S.: Smart materials and structures—IOPscience. *Iopscience.iop.org* (2018). <http://iopscience.iop.org/journal/0964-1726>, Accessed 07 Oct 2018
10. <https://www.popsci.com/article/science/under-light-chameleon-material-changes-color-and-shape>
11. BBC—GCSE bitesize: smart materials. *Bbc.co.uk* (2018). <http://www.bbc.co.uk/schools/gcsebitesize/design/electronics/materialsrev5.shtml>, Accessed 07 Oct 2018
12. Hu, J.: Adaptive and functional polymers, textiles and their applications. Google Books (2018). https://books.google.com/books/about/Adaptive_and_Functional_Polymers_Textile.html?id=YqcsvDMzvcC, Accessed 07 Oct 2018
13. Zhang, M., Wang, X., Sun, Q.: New type intellectualized clothing materials-color-changing materials. *Adv. Mater. Res.* **760–762**, 785–788 (2013)
14. https://in.images.search.yahoo.com/yhs/search;_ylt=Awrxyvli_flboDsA7ujnHgx.;_ylu=X3odMTByYmJwODBkBGNvbG8Dc2czBHBvcwMxBHZ0aWQDBHNIYwNzYw-?p=colour+changing+clothing+in+fashion+industry&fr=yhs-iry-fullyhosted_003&hspart=iry&hsimp=yhs-fullyhosted_003

An Adaptive Framework of Learner Model Using Learner Characteristics for Intelligent Tutoring Systems



Amit Kumar and Neelu Jyoti Ahuja

Abstract Learner Model is the base for providing the adaptivity in the Intelligent Tutoring Systems (ITS) as learner performance data and information are stored in the learner model. Recently, there has been a rapid progress in education delivery through web due to the advancement in internet technology. With the limitations such as lack of adaptive support and presentations reasoned that research has been expanded in the domain of ITS. The aim of the paper is to present an adaptive framework of the learner model using learner characteristics that helps to provide the adaptive presentation and feedback to the prospective learner. Adaptive learner model has three component such as Learner Characteristics Model, Learner Classification Model, and Learner Adaptation Model. Learner Characteristics Model includes the characteristics of learner such as learning style, knowledge levels, and cognitive and meta-cognitive skills. Learner Classification Model classifies the learner into groups based on his/her learning style, levels of knowledge, and performance data and implemented through artificial intelligence technique. The Learner Adaptation Model recommends the tutoring strategy which best suits the learner to provide the adaptation.

Keywords Learner model · Intelligent tutoring system (ITS) · Adaptivity · e-learning · Learning style

1 Introduction

With the growth and advancement in internet technology, there has been rapid progress in education delivery through the web. This mode popularly termed as e-

A. Kumar (✉)

School of Computer Science and Engineering, University of Petroleum and Energy Studies, Dehradun 248001, Uttarakhand, India
e-mail: amitanit007@gmail.com

N. J. Ahuja

University of Petroleum and Energy Studies, Dehradun 248001, Uttarakhand, India
e-mail: neelu@ddn.upes.ac.in

learning, is not a mere delivery of the contents on the web to the prospective learner, but caters to the needs of instructors, and also students, looking for developing their own subject specific repository. E-learning offers education to different learners or gathering of learners whenever and wherever, without any particular emphasis on exploring their learning preferences, needs, and level of knowledge. Each individual characteristics is different from the others and hence the learning process followed by each individual is significantly different from that of the other individual. Learners are different and they have different learning characteristics and can be classified based on his/her characteristics which affect their learning process [1]. This, to a greater degree determines their learning preferences, needs, level of learning, cognitive and meta-cognitive skills. Several studies have been conducted around these differences with an aim to improve the effectiveness of learning while designing the teaching and learning environment [2, 3]. The capacity to customize and adapt to the requirements of an assortment of learners, and to adjust to their diverse background, aptitudes, and abilities to learn is an important part of Intelligent Tutoring System (ITS) [4].

ITS refers to the computerized educational system incorporating the artificial intelligence techniques with minimal intervention of human by adapting the learner needs and preferences. ITS comprises of four main component which are A Learner Model, Tutoring Model, Domain Model, and the Learner Interface Model. However, current tutoring system has several limitations such as to provide an adaptive content and flexible support to the prospective learner [5]. That is the reason, lot of research in the field of computerized educational system/e-learning have taken interest that will suitable for learners who are from different learning backgrounds, learning style, and level of knowledge [6]. It is very difficult and challenging to meet all requirements and to provide the adaptive contents for heterogeneous learners [7]. Therefore, learner modeling is the solution to provide adaptivity and has been introduced in ITS.

Learner modeling is the base for providing the adaptability and personalization in the ITS [8]. Learner modeling was started in 1980 with the main work by Allen, Cohen, and much more Educationalist [9]. In order to automatically adapt each individual learner, ITS needs to monitor learner conduct and activities, and break them down to find how the learner attributes develop after some time [10]. With a specific end goal to develop learner model, it must be considered what learner characteristics are to be modeled and how to model them, ensuring to keep the system up to date [11]. The learner model is developed using the characteristics learning style and preferences, learner knowledge level, affective features such as emotion recognition, cognitive aspect, and meta-cognitive aspects. Hybrid learning style detection mechanism that can help in developing the adaptive tutoring system [10]. Considering the educational games [12], defined that learner modeling can be explicit and implicit. The learner model is mostly instated by either default value or data acquired by administering a questionnaire, thereafter maintained by the system although it can be periodically edited or reviewed by the learner. The problem while initializing a learner model is generally known as a ‘cold-start-problem’ and can be solved by the classification of learner is defined in [13]. In this technique, a learner is assigned into one of the predefined groups based on his/her characteristics and

this enables the system to start quickly, starting by the customized instruction to the learner, specific to the group.

2 Related Work

The related work is centered around the current tutoring system/computerized educational system that implements the learner model using learner characteristics through learner modeling techniques with Artificial Intelligent (AI) approaches [1, 4]. The summary of learner characteristics and modeling techniques used in intelligent tutoring systems/educational hypermedia system/e-learning system has been presented in Table 1.

3 Learner Model

Learner model holds learner information to support the adaptation of instructional material in any personalized learning environment. It holds the learner information and the performance data during learning. The learner model is intended to identify individuality of learner based on the learner characteristics. This section presents an adaptive architecture of the learner model as shown in Fig. 1.

3.1 Adaptive Framework of Learner Model

The learner model architecture is composed of three main components—Learner Characteristics Model, Learner Classification Model, and Learner Adaptation Model

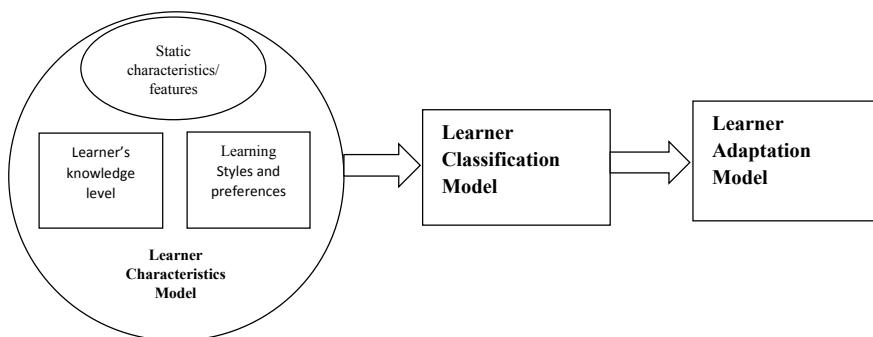


Fig. 1 Architecture of adaptive learner model

Table 1 Summary of learner characteristics and modeling techniques used in ITS/e-learning systems

S.No	Tutoring system	Year	Domain/Subject	Learner characteristics/Features	Learner modeling techniques
1	ICICLE [14]	2000	English	Knowledge level	Overlay, stereotypes
2	MEDEA [15]	2005	OXML	Knowledge level	Overlay
3	VIRGE [16]	2003	English	Knowledge, affective features	Cognitive theories
4	F-SMILE [17]	2002	Electrodynamics	Prior knowledge, errors	Stereotypes, cognitive models
5	SQLT-Web [18]	2003	SQL	Knowledge	Constraints based
6	INFOMAP [19]	2001	Ontology	Prior knowledge, error, and misconception	Overlay, perturbation
7	Web-EasyMath [20]	1999	Algebra	Knowledge level	Stereotypes, machine learning
8	ANDES [21]	2005	Physics	Knowledge, errors, cognitive, meta-cognitive	Constraints based, bayesian networks
9	LeCo-EAD [22]	2004	Computer science	Knowledge, errors and misconception	Perturbation
10	AMPLIA [23]	2008	Medical diagnosis	Knowledge, cognitive	Cognitive theories, bayesian networks
11	E-Teacher	2008	Oops and UML diagrams	Knowledge, learning style	Bayesian network
12	J-LATTE [24]	2009	Java	Knowledge, Errors and misconception	Constraints based modeling

It is proposed to conduct a rigorous review of learner characteristics primarily, Learner knowledge level, learner learning styles, and cognitive abilities, and identify 2–3 sub-characteristics under each of them, and use them for learner classification. Thereafter, an adaptation learner model will be designed that incorporates learner classification as its input and utilizes it to develop an adaptable model to adapt learning content, to offer to the learner. It is proposed that the learning material so offered to the learner is likely to be very closely adapted to the learner, as the learner has been studied and classified in the light of learner characteristics in the current scope of work. Figure 2 presents a framework of the learner model in order to enhance the adaptation and personalization.

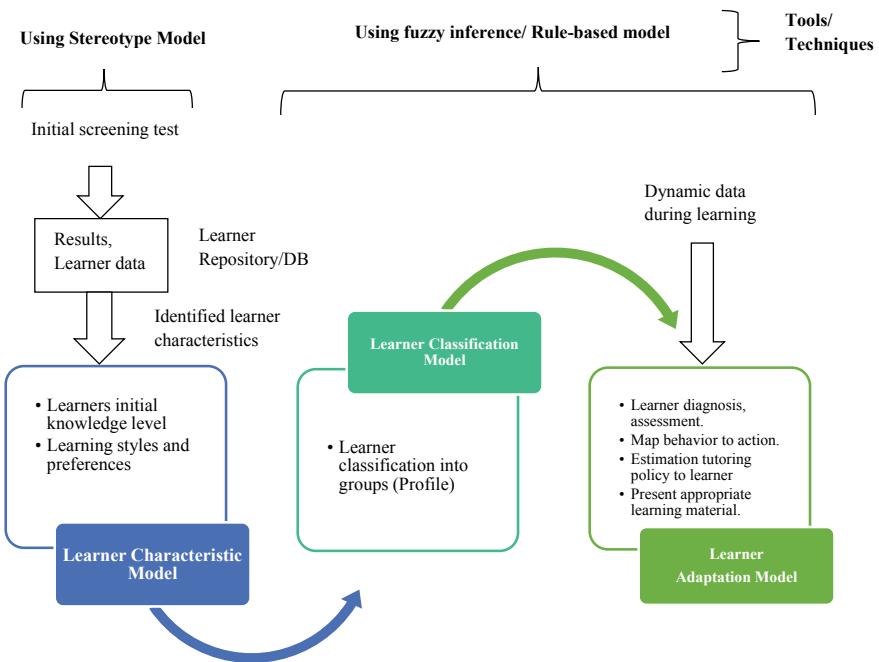


Fig. 2 Components and flow of adaptive learner model

4 Methodology

The methodology of development of an adaptable ITS through the development of a learner model adapted through learner characteristics is presented in this section. This section presents the components and flow of adaptive learner model framework to construct an adaptable learner model based on the learner characteristics shown in Fig. 2.

4.1 Components and Flow of Adaptive Learner Model

4.1.1 Learner Characteristics Model

The focus of this model is on learner characteristics to the model learner to implement adaptation and personalization in the proposed system. Personalization parameter helps to improve the effectiveness of learning and provides flexible support to the learner. Learner characteristics are categorized as static and dynamic features [25]. Static features remain unchanged throughout the learning session and are set before the tutoring session starts. Some examples are name, age, email id, etc. On the other

hand, dynamic features are those that undergo change during the learning session. They are gathered and regularly updated on the basis of learner action or interaction with the system [10]. Investigation and study of existing intelligent tutoring systems and systematic literature review will be undertaken, to explore the different learner characteristics and modeling techniques that have been used in recent past. Development of learner characteristics model is proposed using some characteristic such as knowledge level, learning styles, and cognitive traits in our characteristic model.

However, each individual has a variance of these characteristics. Hence to provide appropriate learning content, and for learner classification, each learner is made to take an initial screening test and accordingly a grade is awarded. The score of each learner under different characteristics is recorded, scaled, and normalized for uniformity. Performance and other details are maintained in learner repository. Tutoring system retrieves and uses for further analysis.

Drilling down the identified learner characteristic, specific sub-characteristics will be identified to model the learner for adaptability of ITS. Various techniques are proposed to be taken up to model these learning characteristics, as shown below [10].

1. Static features using Stereotypes Model
2. Dynamic features using Neural Network Model or Machine Learning Approach

The data of learner characteristics during ongoing learning session is maintained within the learner history, to be used for learner modeling and adaptation.

4.1.2 Learner Classification Model

Using the output of learner characteristics model, the learner classification model is proposed to classify the learners. The objective is to develop learner's groups, which can be allotted to a specific learner, as per learner's judgment and analysis.

4.1.3 Learner Adaptation Model or Diagnostics Model

This model focuses on accurate learner diagnosis in order to assess the learner needs and preferences to adapt the learning material assisting in regulating each individual learner learning pace [26]. The learner assessment process (diagnosis) is proposed to be implemented through reinforcement learning. Reinforcement learning approach is a machine learning technique, in which the system reorganizes itself in a manner to maximize enumerative reward and yield effective results.

It focuses on "what to do" to map the behavior to action. As per the techniques, the learner does not indicate which action to take, the system will automatically discover which action to take based on the learner history or data collected during learning. It means that the adaptable model using reinforcement learning mimics the behavior of an environment. Adaptation module generates an accurate estimation of learner's static and dynamic features to present appropriate learning material.

It consults the learner characteristic model to assess the knowledge level, learning style, and the cognitive traits of learner. Also, the learner's meta-cognitive skills like self-assessment, self-learning, and self-regulation are utilized. The adaptation model seeks this data from learner characteristic and learner classification model.

The purpose is recommending the effective instructional strategy. These learner characteristics assist the adaptive/personalized tutoring system to present appropriate learning contents ensuring learning gain and effectiveness of the proposed system. The expected outcome is that the system can offer task of appropriate complexity and learning the material to the learner leading to improved learning, and assist the learner to become a good learner [26].

5 Conclusion

The research was focused to implement an adaptive framework of learner model using learner characteristics in ITS. Specifically, we accumulated data from different works in the field and distinguished the most vital techniques utilized. The adaptive learner model has been presented with learner characteristics such as learning style, levels of knowledge, and cognitive skills. Based on the characteristics learner model comprises of three sub-models such as learner characteristics model, learner classification model, and learner adaptation model. Stereotype modeling approach has been used to implement the learner characteristics model and learner classification model has been implemented through fuzzy logic modeling technique to classify the learner into groups based on his/her learner characteristics. Learner adaptation model has been implemented through reinforcement learning to assess the learner performance to recommend the tutoring strategy which best suits the learner as per his/her needs and preferences.

Acknowledgements This work is being carried out at University of Petroleum and Energy Studies (UPES) with the reference number SR/CSI/140/2013. The authors thankfully acknowledge the funding support received from Cognitive Science Research Initiative, Department of Science and Technology for the project. The Authors thank the management of UPES, for supporting the work and granting permission to publish it.

References

1. Kumar, A., Singh, N., Ahuja, N.J.: Learner characteristics based learning style models classification and its implications on teaching. *Int. J. Pure Appl. Math. (IJPAM)* **118**(20), 175–184 (2017)
2. Bozkurt, O., Aydoğdu, M.: A comparative analysis of the effect of dunn and dunn learning styles model and traditional teaching method on 6th grade students' achievement levels and attitudes in science education lesson. *Elem. Educ. Online* **8**(3), 741–754 (2009)
3. Demirbaş, O.O., Demirkan, H.: Focus on architectural design process through learning styles. *Des. Stud.* **24**(5), 437–456 (2003)

4. Kumar, A., Singh, N., Ahuja, N.J.: Learning styles based adaptive intelligent tutoring systems: document analysis of articles published between 2001 and 2016. *Int. J. Cogn. Res. Sci. Eng. Educ.* (IJCRSEE) **5**(2), 83–98 (2017)
5. Xu, D., Wang, H., Su, K.: Intelligent student profiling with fuzzy models. In: Proceedings of the 35th Annual Hawaii International Conference on System Sciences, pp. 8. IEEE (2002)
6. Schiaffino, S., Garcia, P., Amaldi, A.: eTeacher: providing personalized assistance to e-learning students. *Comput. Educ.* **51**(4), 1744–1754 (2008)
7. Lo, J.J., Chan, Y.C., Yeh, S.W.: Designing an adaptive web-based learning system based on students' cognitive styles identified online. *Comput. Educ.* **58**(1), 209–222 (2012)
8. Ayala, A.P., Sossa, H.: Semantic representation and management of student models: an approach to adapt lecture sequencing to enhance learning. In: Mexican International Conference on Artificial Intelligence, pp. 175–186. Springer, Berlin, Heidelberg (2010)
9. Kobsa, A.: Generic user modeling systems. *User Model. User-Adap. Interact.* **11**(1–2), 49–63 (2001)
10. Jeremic, Z., Jovanovic, J., Gasevic, D.: Student modeling and assessment in intelligent tutoring of software patterns. *Expert Syst. Appl.* **39**(1), 210–222 (2012)
11. Mourlas, C., (ed.): Cognitive and Emotional Processes in Web-Based Education: Integrating Human Factors and Personalization: integrating Human Factors and Personalization. IGI Global (2009)
12. Khenissi, M.A., Essalmi, F., Jemni, M., et al.: Learner modeling using educational games: a review of the literature. *Smart Learn. Environ.* (2015)
13. Denaux, R., Dimitrova, V., Aroyo, L.: Integrating open user modeling and learning content management for the semantic web. In: International Conference on User Modeling, pp. 9–18. Springer, Berlin, Heidelberg (2005)
14. Michaud, L.N., McCoy, K.F., Pennington, C.A.: An intelligent tutoring system for deaf learners of written English. In: Proceedings of the Fourth International ACM Conference on Assistive Technologies, pp. 92–100. ACM (2000)
15. Trella, M., Carmona, C., Conejo, R.: MEDEA: an open service-based learning platform for developing intelligent educational systems for the web. In: Workshop on Adaptive Systems for Web-Based Education: tools and Reusability (AIED'05), pp. 27–34 (2005)
16. Virvou, M., Katsionis, G.: VIRGE: tutoring English over the web through a game. In: The 3rd IEEE International Conference on Advanced Learning Technologies, 2003. Proceedings, p. 469. IEEE (2003)
17. Virvou, M., Kabassi, K.: F-SMILE: An intelligent multi-agent learning environment. In: Proceedings of 2002 IEEE International Conference on Advanced Learning Technologies-ICALT (2002)
18. Mitrovic, A.: An intelligent SQL tutor on the web. *Int. J. Artif. Intell. Educ.* **13**(2–4), 173–197 (2003)
19. Hsu, W.L., Wu, S.H., Chen, Y.S.: Event identification based on the information map-INFOMAP. In: 2001 IEEE International Conference on Systems, Man, and Cybernetics, vol. 3, pp. 1661–1666. IEEE (2001)
20. Virvou, M., Tsiriga, V.: EasyMath: a multimedia tutoring system for algebra. In: EdMedia: world Conference on Educational Media and Technology. Association for the Advancement of Computing in Education (AACE), pp. 933–838 (1999)
21. Vanlehn, K., Lynch, C., Schulze, K., Shapiro, J.A., Shelby, R., Taylor, L., Wintersgill, M.: The andes physics tutoring system: lessons learned. *Int. J. Artif. Intell. Educ.* **15**(3), 147–204 (2005)
22. Faraco, R.A., Rosatelli, M.C., Gauthier, F.A.: A learning companion system for distance education in computer science. In: Anais do XXIV Congresso da Sociedade Brasileira de Computação (SBC 2004), XII Workshop de Educação em Computação (WEI 2004) (2004)
23. Vicari, R., Flores, C.D., Seixas, L., Gluz, J.C., Coelho, H.: AMPLIA: a probabilistic learning environment. *Int. J. Artif. Intell. Educ.* **18**(4), 347–373 (2008)
24. Holland, J., Mitrovic, A., Martin, B.: J-LATTE: a constraint-based Tutor for Java (2009)
25. Essalmi, F., Ayed, L.J.B., Jemni, M., Graf, S.: A fully personalization strategy of e-learning scenarios. *Comput. Hum. Behav.* **26**(4), 581–591 (2010)

26. Dorça, F.A., Lima, L.V., Fernandes, M.A., Lopes, C.R.: Comparing strategies for modeling students learning styles through reinforcement learning in adaptive and intelligent educational systems: an experimental analysis. *Expert Syst. Appl.* **40**(6), 2092–2101 (2013)

A Study of Solar Energy in India; Utility, Status, and Procurement



Kamal Kashyap, S. M. Towhidul Islam Nayim, Robin Thakur, Anil Kumar and Sashank Thapa

Abstract Energy consumption is a global issue. Most of the countries are struggling to fulfill their energy needs due to their dependency on nonrenewable energy resources. As nonrenewable energy resources are depleting, this is very important to find out new possibilities in renewable energy. India with a total of 300 sunny days has a lot of potential in solar energy. This research paper gives an overview of the current scenario and the future possibilities of solar power in India. The research paper also focuses on different aspects related to solar power and initiatives taken by the government of India to fulfill the countries future energy needs. This study also may help to make decisions related to solar energy sector in India.

Keywords Renewable energy · Solar power · Energy potential

1 Introduction

With the vision of *Electricity to Every One*, India is finding out new possibilities for energy generation in the field of renewable energy. With the total energy consumption of 1,068,923 MU and production of 1,030,785 MU India is lacking 38,138 MU to fulfill its energy need. The dependency on coal in case of energy generation is very

K. Kashyap · S. M. Towhidul Islam Nayim

Department of Mechanical Engineering, A P Goyal Shimla University, Shimla, HP, India
e-mail: kamal.maximus23@gmail.com

S. M. Towhidul Islam Nayim

e-mail: islam.naayim@gmail.com

R. Thakur (✉) · A. Kumar · S. Thapa

School of Mechanical & Civil Engineering, Shoolini University, Solan, HP, India
e-mail: robinsampark.thakur@gmail.com

A. Kumar

e-mail: anilkumar88242@gmail.com

S. Thapa

e-mail: mech12400@gmail.com

high as 56.47% of the total energy demand is fulfilled with coal, the left 44.53% is, Hydro with the total percentage of below 10%, oil with 2.18%, renewable energy with 2.18% and nuclear with the minimum dependency with just 1.32%. Solar power with the aggregate production of just 23 GW can be the major source of energy in near future as the Gov. with the target of 22 GW grid connected solar projects by 2022 has been increased to 100 GW.

2 State-Wise Renewable Energy Potential

India has estimated renewable energy potential through out the country. To achieve the sound and healthy amount of renewable energy source, government and non-government organizations already have initiated the steps for achieving the target that has been set for the next decades. As compared to source wise power generation capacity, 57244.23 MW renewable energy is installed whereas only from coal 192162.88 MW is installed throughout India till 2017 [1]. Along with all India installed capacity, every state has different conditions in the potential of renewable energy. Rajasthan, Jammu & Kashmir, Maharashtra, Madhya Pradesh are being ahead among the other states of India as shown in Fig. 1.

Most promising states potential are being shown in the above graphical representation. From there Telangana is the state where only solar energy is developed where the rest of the others also have opportunity to build up solar energy [2].

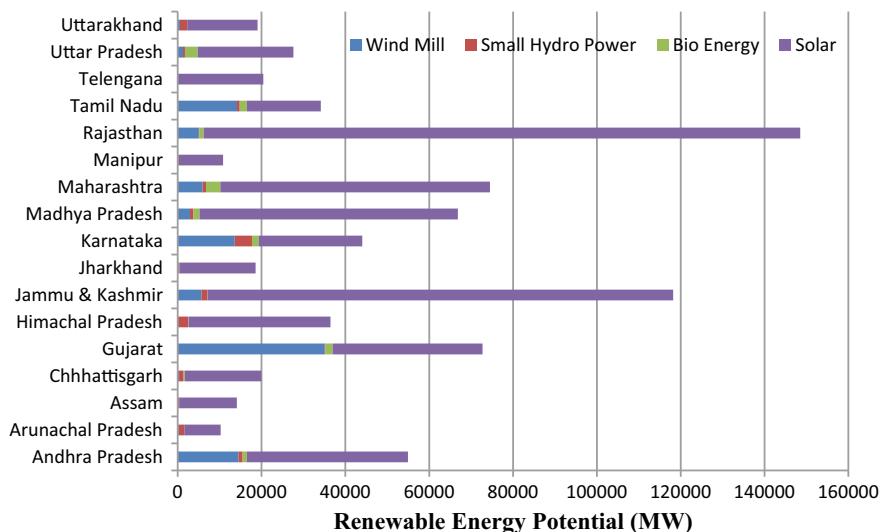


Fig. 1 States having the maximum potential of renewable energy in India. *Source* Annual Report 2016–2017 by MNRE, Government of India

3 Solar Radiation Zone

India lies in the sunny belt of the world. As most parts of India are under 300 days of sunshine and solar energy daily average varies from 4 to 7 KWh/m² depending upon location [3]. About 5000 kWh/Year of power is received by the radiation over Indian land area. And most of the areas are receiving 4–7 kWh/m² power in every day [4]. Rajasthan is the maximum solar radiation getting area (5.8–6.6 Kwh/m²). Northern Gujarat, part of Andhra Pradesh, Madhya Pradesh, part of Ladakh region are also under a large radiation which near about 5.8–6.2 Kwh/m² [5]. Most promising zone (state) under solar radiation is shown in Fig. 2.

4 Scaling Up Renewable Energy for 2022

175 GW target of renewable energy capacity is to be designed by the Government of India. Where solar energy is given maximum importance and has been scaled up 100 GW capacity by the year 2022. This target is compromised into two categories, 40 GW for rooftop and 60 GW for large and medium grid-connected solar power project. Through this India will become one of the largest Green Energy producers in the world [6]. Target by 2022 is compared (in percentage) in Fig. 3.

Fig. 2 Maximum solar radiation zone (considering states)



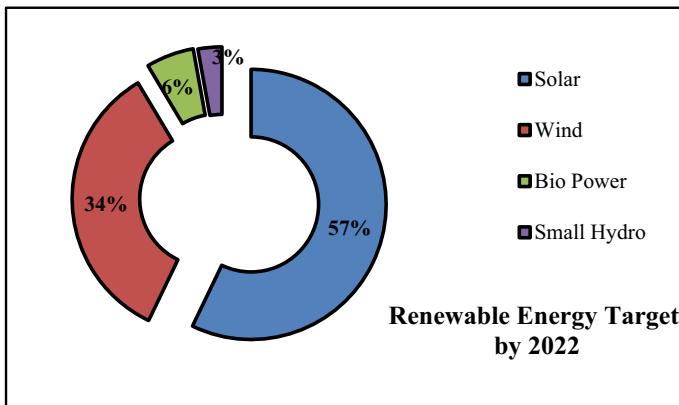


Fig. 3 Renewable energy capacity target by 2022. *Source* Annual Report 2016–2017 by Government of India, Ministry of New and Renewable Energy

In terms of GW capacity, 60 GW target sets for wind, 10 GW for bio power, and 5 GW for small hydro [7]. As per the target 100,000 MW Grid connected cumulative target is set out by the year 2021–22 under National Solar Mission [8].

5 Rooftop Solar Power Utilization

Developed countries are trying to encourage more rooftop solar energy growth. Although in India the utility scale and rooftop solar ratio is not satisfying enough. It is near about 10–12% share of overall solar capacity in the country which is lower than other. Like Australia is using 97% share of overall solar capacity, China is using 18%, USA is using 46%, and Germany is using 73% [9].

6 Status of Grid Connected Solar Projects

Till March 2017 some capacity was commissioned with total cumulative capacity. So the total cumulative capacity is changed in some states. Total cumulative capacity is reached in last year till March, 2017; for Andhra Pradesh 1867.23 MW, for Bihar 108 MW, for Chhattisgarh 128.86 MW, for Delhi 40.27 MW, for Gujarat 1249.37 MW, for Haryana 81.4 MW, for Karnataka 1027.84 MW, for Kerala 74.2 MW, for Madhya Pradesh 857.04 MW, for Maharashtra 452 MW, for Odisha 79.42 MW, for Punjab 793.95 MW, for Rajasthan 1812.93 MW, for Tamil Nadu 1691.83 MW, for Telangana 1286.98 MW, for Uttar Pradesh 336.73 MW, for Uttarakhand 233.49 MW and for rest of the others 166.28 MW. New capacity is also being

added; for Andhra Pradesh 271 MW, for Bihar 33 MW, for Chhattisgarh 0.05 MW, for Delhi 16.95 MW, for Gujarat 41.81 MW, for Haryana 110.04 MW, for Karnataka 464.54 MW, for Kerala 14 MW, for Madhya Pradesh 282.95 MW, for Maharashtra 62.64 MW, for Odisha 0.07 MW, for Punjab 82.85 MW, for Rajasthan 433.55 MW, for Tamil Nadu 20.24 MW, for Telangana 1283.43 MW, for Uttar Pradesh 171.01 MW, for Uttarakhand 13.4 MW, and for rest of the others 13.73 MW. These total cumulative amounts and the new commissioned amount are being added for introducing the status of grid connected solar projects till October 2017. In total the final cumulative capacity till October 2017 is reached 15604.76 MW where newly commissioned capacity has the amount of 3315.92 MW [10].

7 Green Cities of India's Plan

World is growing rapidly. With the development of economy and urbanization, the demand for energy is increasing day by day. Nowadays every growing city is planning for reducing the emission of GHS to prevent the effect of Climate change. For this purpose, world is introduced with Solar Cities. It is nothing but to reduce the dependency in conventional energy and to increase the dependency on renewable energy. Australia, USA, Germany, and many more countries like these are utilizing the solar energy [9]. Government of India and electricity utilities are finding many

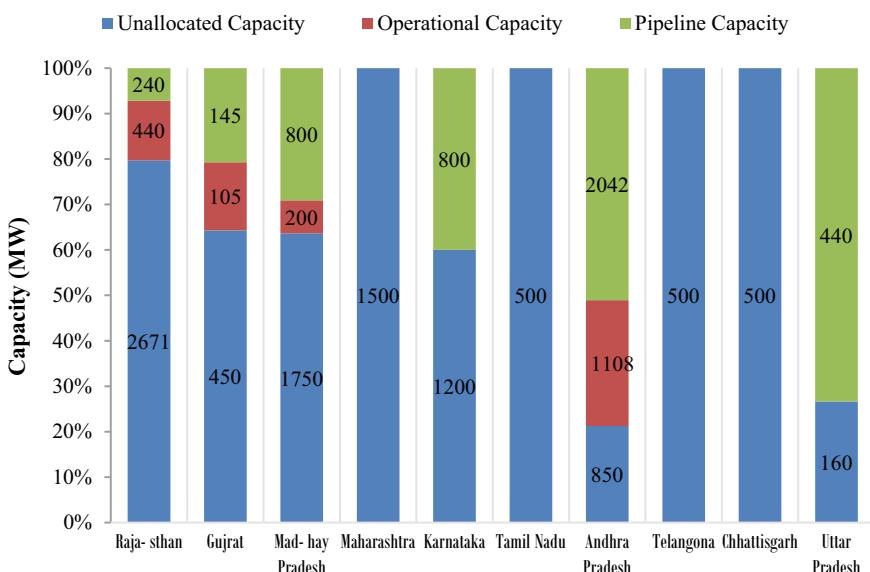


Fig. 4 The implementation status of solar cities in different states. *Source* India Solar Handbook 2017, © 2017 BRIDGE TO INDIA Energy Private Limited

difficulties to cope with the demand. With this view “Development of Solar Cities” programme is designed to transform the cities into Solar Cities or Green Cities. The aims of these solar cities are to reduce the demand of conventional energy at least 10% [11]. 60 cities are introduced in the design of Solar Cities during the 11th plan period where minimum one city and maximum five cities in a state will be supported by the ministry. So far fulfillments of criteria some cities are identified and given principle approval 48 cities are selected. Among these cities 31 cities are received principle approvals [12]. Some prominent states are scaled in Fig. 4.

The highest working progress is achieved in Uttar Pradesh. Where Maharashtra, Tamil Naidu, Telangana, Chhattisgarh are in the least. But in amount the highest working progress is done in Madhya Pradesh about 1000 MW. In Maharashtra total capacities are unallocated. The highest working progress is seen in Andhra Pradesh, about 80% of total capacity under progress. For Rajasthan 2671 MW is unallocated and 20% of total capacity is under progress [13].

8 Sector-Wise Indian Market of Solar

Indian solar market is increasing day by day. Government simplifies their policies. Lots of projects are running, also a huge number of projects are in the pipeline. As the market grows, also project sizes are increasing, domestic as well as international developers are already get through with this big sector. As for their contributions already are in great manner, their footprint is marked. According to different areas capacity in MW are graphically categorized. Status of the capacity by different sectors is presented as commissioned capacity versus capacity under development in Fig. 5.

5725 MW and 5414 MW are respectively presenting the condition of Indian corporates in commissioned capacity and sector under development. Under public

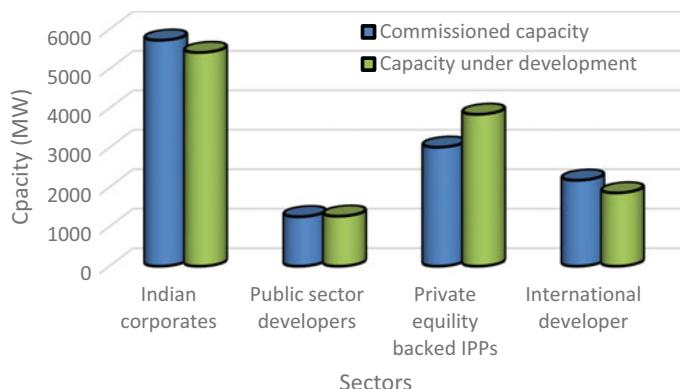


Fig. 5 Project development sharing. *Source India Solar Handbook 2017, © 2017 BRIDGE TO INDIA Energy Private Limited*

sectors commissioned capacity is 1252 MW and capacity under development is 1265 MW. Under private (IPPs) sector is developed as for commissioned capacity 3012 MW and for capacity under development 3847 MW. International developers are also marked significantly in the development of solar sectors and where commissioned capacity is 2173 MW and capacity under development is 1855 MW [14].

9 Government Acts and Mission

In January 11, 2009 Jawaharlal Nehru National Solar Mission (JNNSM) was launched and was committed to set a target of 20,000 MW by 2022 for Grid Connected Solar Projects with three phases of approach [15, 16]. Indian Government has come out with acts as well as their policies. The Electricity Act 2003, National Electricity Policy 2005, Tariff Policy 2005 can clarify in different manners. By which the stand of government against solar energy and the stats as well as future aspects are introduced. The share of electricity from nonconventional resources would need to be increased such purchase by distribution companies shall be through competitive process' comes under National Electricity Policy 2005 [17]. Under Sect. 63 of Electricity act 2003 states that- "The Appropriate Commission shall adopt the tariff if such tariff has been determined through transparent process of bidding in accordance with the guidelines issued by the Central Government" [18]. On the other hand in Tariff Policy 2006, it states that Appropriate Commission shall decide a minimum percentage for purchase of energy from nonconventional source according to the availability of resources in that region and its impact on retail tariffs [19, 20].

10 Advancement in Renewable Energy Field

India is ranked fifth in the electricity generation in the world [21] but most of the electricity generation comes from coal or thermal energy. And these thermal plants emit a large number of toxic gases. But India is situated under sunny belt region with a good amount of radiation on its land. Government utilizes this opportunity and Jawaharlal Nehru National Solar Mission (JNNSM) is launched in 2009. The total investment will be around Rs. 6,00,000 cr. (@ Rs. 6 cr. per MW at present rate) for 100 GW power generation under this mission [22]. As a country, India was the first who creates the Ministry of NonConventional Energy Resources in 1980 [21]. Though India has great utility, but there are also barriers . Storage problem, more investment requirements, lands settlement are the areas of difficulties. And the statics shows that the working progress for every state is not linear, some states are growing very fast some are growing low. This nonlinearity has to be minimized. Also the amount of work in private sectors are increasing more than the public sectors. Nowadays international companies are also investing in the field of solar as India is promising to strong in renewable energy. Another very important aspect is

environment-friendly power generation system, under new solar target of 100 GW is expected to abate over 170 million tons of CO₂ over its life cycle [22]. It has to be admired that India is ranked 11 in solar power generation in the world as on January 2014 [23].

11 Conclusion

This paper starts with the section Solar Potential and finished with Government Acts. It reflects that this paper is continued with a manner which gradually comes from the source to the use. It also contains the market of solar in India as well as market comparison with other developed countries. From this study it is clear that how the solar power is increasing, and how much capital is used in this sector. Government steps, their policies are introduced here. Targets for upcoming years also are presented. The contribution of public sector, private sector also be understandable. Even it is possible to have idea about manufacturer's market condition from this paper. The success and the achievements in solar sector also reflect in this paper. By this complete study it is sure that India has a great opportunity to build solar strength. From the last couple of years, the progress is remarkable.

References

1. http://www.cea.nic.in/reports/monthly/installcapacity/2017/install_capacity-03.pdf
2. Renewable Energy Potential (section 1.7), Annual Report 2016–2017 by Government of India, Ministry of New and Renewable Energy
3. Renewable energy in India. MNES report, Government of India (2001)
4. Amita, U., Soni, M.S.: Concentrating solar power – Technology, potential and policy in India. Renew. Sustain. Energy Rev. **15**, 5161–5175 (2011)
5. Garud, S., Purohit, I.: Making solar thermal power generation in India a reality—Overview of technologies, opportunities and challenges. In: The Energy and Resources Institute (TERI), Darbari Seth Block, IHC Complex, Lodhi Road, New Delhi 110003, India
6. RENEWABLE ENERGY TARGETS (sections 1.12), Annual Report 2016–2017 by Government of India, Ministry of New and Renewable Energy
7. Press Information Bureau, Government of India, Ministry of New and Renewable Energy; 19 July 2018 15:45 IST
8. Jawaharlal Nehru National Solar Mission, 23-July-2015 16:21 IST; Press Information Bureau, Government of India, Ministry of New and Renewable Energy
9. International market review, India Solar Handbook 2017, © 2017 BRIDGE TO INDIA Energy Private Limited
10. Ministry of New and Renewable Energy, Government of India; March, 2017
11. <https://mnre.gov.in/solar-cities>
12. Cities identified for developing as Solar Cities (section 8) & Cities Sanctioned (section 9), Green Cities; Ministry of New and Renewable Energy, Government of India
13. Development of solar parks, India Solar Handbook 2017, © 2017 BRIDGE TO INDIA Energy Private Limited

14. Project development landscape, India Solar Handbook 2017, © 2017 BRIDGE TO INDIA Energy Private Limited
15. LollaSand Roy, S.B.: Wind and Solar Resources in India. *Energy Procedia*. **70**, 187–192 (2015)
16. Ojha, A.K., Gaur, G.K.: Solar energy and economic development in India: a review. *Int. J. Emerg. Technol. Adv. Eng.* **4**, 184–189 (2014)
17. India Solar Handbook: Bridge to India. Bridge to India Energy Pvt. Ltd, India (2015)
18. Solar Energy in India. Solar Energy Association, Tamil Nadu
19. Ishan, P., Pallav, P.: Techno-economic evaluation of concentrating solar power generation in India. *Energy Policy*. **38**, 3015–3039 (2010)
20. Dubey, S., Chamoli, S.: Indian scenario of solar energy and its application in cooling systems: a review. *Int. J. Eng. Res. Technol.* **6**, 571–578 (2013)
21. Load Generation and Balance Report, Central Electricity Authority, Ministry of Power, Government of India. Central Electricity Authority. 2015–16
22. Sharma, B.D.: Performance of Solar Power Plants in India. Central Electricity Regulatory Commission New Delhi (2011)
23. Vikas, K., et al.: Status of solar wind renewable energy in India. *Renew. Sustain. Energy Rev.* **27**, 1–10 (2013)

Necessary Precautions in Cognitive Tutoring System



**Kevin Vora, Shashvat Shah, Harshad Harsoda, Jeel Sheth
and Ankit Thakkar**

Abstract Systems that enhance the cognitive capability of the user can be used for the educational purpose. Physiological and psychological data are the key to figure out the human cognitive states. Using electroencephalography (EEG), heart rate and webcam a cognitive tutoring system is proposed in this paper with an objective to give feedback to the users to pay attention if the user is not attentive while video session is being played. Necessary precautions are suggested in the paper by analyzing feedback received from the user that helps to understand various aspects to be considered while designing such a system.

Keywords Electroencephalography · Tutoring system · Cognitive computing · Heart rate sensor

1 Introduction

Valiant defined cognitive computing as “a discipline that links together neurobiology, cognitive psychology and artificial intelligence” [1]. However, there is no commonly accepted definition of cognition [2], which makes it difficult to define a cognitive computing system. In this paper, designing of a cognitive tutoring system is discussed

K. Vora (✉) · S. Shah · H. Harsoda · J. Sheth · A. Thakkar
Department of Computer Science and Engineering, Institute of Technology,
Nirma University, Ahmedabad 382 481, Gujarat, India
e-mail: 15BIT066@nirmauni.ac.in

S. Shah
e-mail: 15BIT054@nirmauni.ac.in

H. Harsoda
e-mail: 15BIT012@nirmauni.ac.in

J. Sheth
e-mail: 15BIT058@nirmauni.ac.in

A. Thakkar
e-mail: 15BIT052@nirmauni.ac.in; ankit.thakkar@nirmauni.ac.in

that provides feedback to the students when they lose their attention while watching a video session on a certain topic. The system is a cognitive tutoring system as the various physiological, and psychological features are used to measure the required attention level of the students. The system gives appropriate feedback during the tutoring phase to enhance the learning of the student. Feedback is taken from each user by considering various parameters related to the usage of the proposed system. The user experience is analyzed with and without the use of a cognitive tutoring system while learning from video session and necessary precautions are suggested that should be carefully handled while designing of the cognitive tutoring system.

The rest of the paper is organized as follows: Literature review is presented in Sect. 2, the proposed approach to design cognitive tutoring system is presented in the Sect. 3, setup, and result discussions are presented in the Sect. 4, and Concluding remarks are given in the Sect. 5.

2 Literature Review

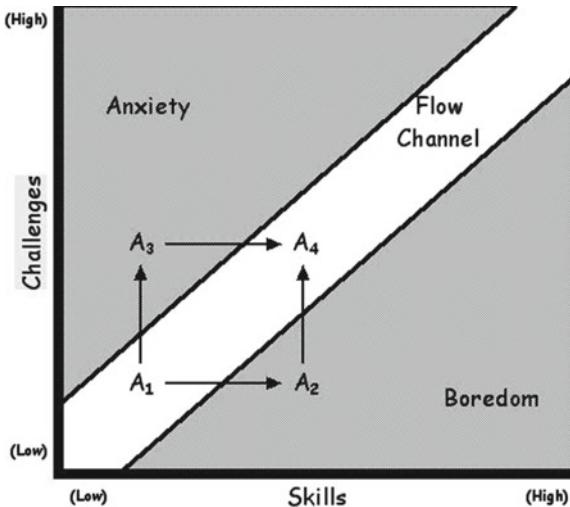
Sensory and biosignals play an important role in determining the emotional state of a person [3]. It is shown in the [4] that the learning experience can be significantly enhanced by effectively understanding the affective state of a user. There are several studies [5, 6] wherein affective states of the users are determined with good accuracy. The dimensions that are usually adopted by the learning theories and models includes arousal (deactivating/activating), valence (negative/positive) intensity (low-intense) duration (short-long), the frequency of its occurrence (seldom-frequent), and time dimension (retrospective like relief, actual like enjoyment, perspective like hope) [7].

By using sensors that can record Electroencephalogram (EEG) signals of the pre-frontal cortex, it is possible to determine the intensity of attention. It has been shown that sensors are more reliable in determining the arousal (deactivating/activating) than valence [8]. Thus, the proposed system uses single-channel EEG headset to obtain the attention level of a user that can help to enhance the learning experience of a user.

Csikszentmihaly [9] defined “The Zone” or “The Flow”, where people are so intensely indulged in the activity, that they lose track of the time. It has been suggested that the learning experience is enhanced when a person spends more time in this flow state. The mapping of the states on a 2D plane is shown in Fig. 1, where X-axis represents the skill level and the Y-axis represents the challenge level, and the flow state lies on the diagonal from the origin in this plane. According to [6], it seems that using the intelligent system and “The Flow Theory” together will help to develop a system that can be leveraged to help people to spend more time in “The Zone” or “The Flow”.

Determining the effective state using an intelligent system is still in its infancy [10]. It is shown that eye movements help to understand visual attention level of a user [11]. Hence, it can be inferred that the amount of time a person spends looking at the screen when a video lecture is being played has a direct relationship with the attention

Fig. 1 Position of flow state with respect to skills and challenges [9]



levels of that person. The study [12] showed that there is a significant relationship between the attention/arousal level and the screen/ message content on the screen. It is also suggested that there is a significant relationship between heart rate variability and the cognitive performance [13]. Thus, a cognitive tutoring system is presented in this paper which is designed using EEG headset (NeuroSky Mindwave headset [14]), a heart rate sensor, and a webcam that aids the human learning process.

3 Proposed Approach

A cognitive tutoring system is designed using three sensors named EEG headset, a heart rate sensor, and a webcam. The use of each of this sensor is as follows:

- The brain activity of a user changes with the time as the user faces different mental attention levels while learning. This can be measured by the EEG headset. The attention level of a user's cognitive state can be quantized to a single value [15]. The user is said to be out of attention if the attention level of the user falls below the threshold T_b .
- The heart rate of a user is an essential physiological parameter to determine the cognitive state of the user [13]. A user is said to be out of the flow channel if the reported heart rate value exceeds the threshold value T_h . The value of T_h is set according to [16].
- To track whether the user is focusing on the screen or not, the system analyzed the user's head position using the face pose angle method. The system pauses the video being played by the user when the system finds that the user is not looking on the screen. This enforces the user to resume the video manually.

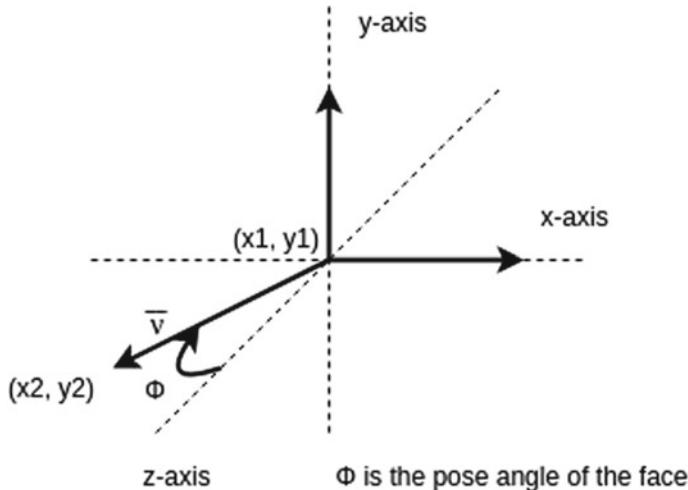


Fig. 2 Estimation of the pose of face

The proposed system uses the method given in [17] to find the pose of the face and the pictorial representation of the same is shown in Fig. 2. The axes represented by the dotted lines indicate the 3D axes for reference, whose XY plane is parallel to the screen itself. The pose of the face obtained by the algorithm given in [17]. The algorithm returns three orthogonal vectors represented by solid lines that intersect at the origin of the axes. The cosine inverse between the vectors \bar{v} and the unit vector lying on the Z -axes (see Fig. 2) represents the pose angle of the face.

The objective of the system is to assist users in their learning task by notifying them whenever their attention falls below the required attention level. The sensors utilized in this experimentation helps to determine the attention levels of the users. The user will be notified by an appropriate message to pay more attention for the learning task if any of the required parameters affecting the learning is not within the specified threshold. Thus, by using the identified sensors to monitor the real-time data of the user, one can measure the attention level of the user and gives the feedback to the user in the form of notification to pay more attention toward video tutoring. This may improve the learning of the user with the video tutoring system.

4 Experimental Setup and Result Discussions

As shown in the Fig. 3 the experimentation is performed using three sensors as discussed in the Sect. 3 wherein a student needs to appear for a test of 10 marks after going through a video session. All the students with the same educational background were selected randomly to participate in the experimentation. The video shown to

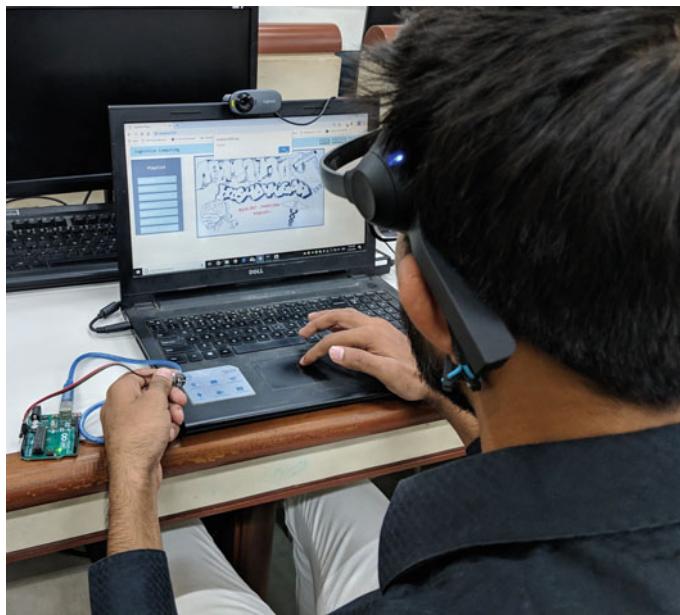


Fig. 3 Experimental setup using webcam, Neurosky EEG headset, and heart rate sensor

the user is from a different field of study. This helps to avoid bias. The students were randomly divided into two groups of the same size:

- Group-1: Students giving the test without the use of cognitive tutoring system.
- Group-2: Students giving the test with the use of cognitive tutoring system.

The results of Group-1 and Group-2 are presented in Tables 1 and 2, respectively. The subject column represents the students participating in the experimentation, Gender column represents the gender of the student, and the score shows the marks obtained in the test out of 10. It is concluded that marks obtained in the presence of cognitive tutoring system were lower than marks obtained without the use of the cognitive tutoring system. The results were also verified statistically using one-tailed Wilcoxon signed rank test [18] and the p-value obtained was less than 0.05.

The feedback received from the students regarding the experimentation for the various parameters named engagement in content, friendly environment, difficulty of the test, negative impact of EEG, negative impact of heart rate sensor, negative impact of webcam, and disturbance due to the pop-up notification. Each parameter to be answered on a scale of 0–5. For the Group-1, the last four parameters were not applicable.

The feedback received from the students of Group-1 and Group-2 is shown in Fig. 4a, b, respectively. It can be concluded from the feedback that the students of Group-1 found the environment to be comfortable, the content of the video was able to engage them, and the test was of average difficulty. However, the feedback of

Table 1 Results of Group-1

Subject	Gender	Score
Student 1	Female	6
Student 2	Female	7
Student 3	Male	7
Student 4	Female	1
Student 5	Male	6
Student 6	Male	7
Student 7	Male	9
Student 8	Female	6
Student 9	Male	5
Student 10	Male	8
Student 11	Male	9
Student 12	Female	7

Table 2 Results of Group-2

Subject	Gender	Score
Student 13	Male	3
Student 14	Male	4
Student 15	Female	4
Student 16	Male	1
Student 17	Male	6
Student 18	Male	1
Student 19	Female	2
Student 20	Female	4
Student 21	Female	2
Student 22	Male	3
Student 23	Male	5
Student 24	Female	1

Group-2 students was opposite to that of Group-1. Plausible reasons for the same could be as follows:

- The continuous notification from the cognitive tutoring system may result in less engagement of content.
- Due to the feeling of being continuously monitored by the devices, students felt restrictive environment. The feedback suggests that the negative impact of the webcam was more compared to the EEG headset and the heart rate sensor. This is because the students were being observed by the webcam.

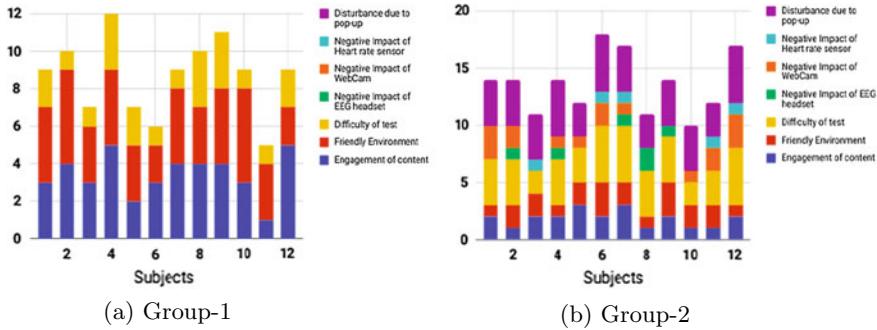


Fig. 4 Feedbacks of two groups

- The student was not able to learn due to disturbance from notification of the system. The average number of notifications were around 11 per minute for the experimentation carried out with the use of the cognitive tutoring system. As a result, the test was difficult for the Group-2 students.

5 Concluding Remarks

Developing a cognitive computing system for intelligent tutoring is an upcoming research problem. Using EEG, heart rate sensor and webcam such a system experimented. By analyzing the results obtained with the use of the system, various precautions like restrictive study environment, disturbance from notifications, and continuous observation of a student, should be carefully handled to avoid the negative impact of the sensors which is a challenging task to design an intelligent tutoring system. However, successful implementation of such a system would help to improve the learning process of students.

Acknowledgements The research is carried out as a part of Idea Lab project funded by the Institute of Technology, Nirma University. The authors would like to thank Nirma University at Gandhinagar (Gujarat, India) for providing us necessary data and permissions to use them in our study. The authors also would like to acknowledge the participants.

References

1. Valiant, L.G.: Cognitive computation. In: Proceedings of the IEEE 54th Annual Symposium on Foundations of Computer Science, pp. 2–2 (1995)
2. Modha, D.S., et al.: Cognitive computing. Commun. ACM **54**(8), 62–71 (2011)
3. Vora, K., Shah, S., Harsoda, H., Sheth, J., Agarwal, S., Thakkar, A., Mankad, S.H.: Motion recognition from sensory and bio-signals: a survey. In: Proceedings of the 2nd International

- Conference on Data Engineering and Communication Technology, pp. 345–355. Springer, Singapore (2019)
- 4. Feidakis, M., Daradoumis, T., Caballe, S.: Emotion measurement in intelligent tutoring systems: what, when and how to measure. In: 3rd IEEE International Conference on Intelligent Networking and Collaborative Systems (2011)
 - 5. Calvo, R.: Incorporating affect into educational design patterns and technologies. In: Proceedings of the 9th IEEE International Conference on Advanced Learning Technologies, Riga, Latvia, 14–18 July 2009
 - 6. Arroyo, I., Cooper, D., Burleson, W., Woolf, B.P., Muldner, K., Christopherson, R.: Emotion sensors go to school. In: Proceedings of International Conference on Artificial Intelligence in Education (AIED), pp. 17–24. IOS Press, UK, Brighton, 6–10th July 2009
 - 7. Hascher, T.: Learning and emotion: perspectives for theory and research. *Euro. Educ. Res. J.* **9**, 13–28 (2010)
 - 8. Zimmermann, P.G.: Beyond usability-measuring aspects of user experience, Doctoral dissertation, Swiss Federal Institute of Technology, Zurich (2008)
 - 9. Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. Harper and Row, New York (1990)
 - 10. Picard, R.W., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D., Machover, T., Resnick, M., Roy, D., Strohecker, C.: Affective learning-a manifesto. *BT Technol. J.* **22**, pp. 253–269 (2004)
 - 11. Duc, A.H., Bays, P., Husain, M.: Eye movements as a probe of attention. *Progress in Brain Research*, vol. 171, pp. 403–411. Elsevier (2008)
 - 12. Reeves, B., Lang, A., Kim, E.Y., Tatar, D.: The effects of screen size and message content on attention and arousal. *Media Psychol.* (1999)
 - 13. Luque-Casado, A., Zabala, M., Morales, E., Mateo-March, M., Sanabria, D.: Cognitive performance and heart rate variability: the influence of fitness level. *PloS one J.* (2013)
 - 14. Electroencephalography (EEG) sensor. <http://neurosky.com/biosensors/eegsensor/>, Last Access 29 Sept 18
 - 15. Liu, N.-H., Chiang, C.-Y., Chu, H.-C.: Recognizing the degree of human attention using EEG signals from mobile sensor (2013)
 - 16. Dimitriev, D.A., Saperova, E.V., Dimitriev, A.D.: State anxiety and nonlinear dynamics of heart rate variability in students. *PLoS ONE* **11**(1), e0146131 (2016). <https://doi.org/10.1371/journal.pone.0146131>
 - 17. Algorithm for calculation. https://github.com/mpatacchiola/deepgaze/blob/master/examples/ex_pnp_head_pose_estimation_webcam.py, Last Access 12 Oct 18
 - 18. Wilcoxon Signed-Ranks Test Calculator. <https://www.socscistatistics.com/tests/signedranks/Default.aspx>, Last Access 4 Oct 18

Node-Level Self-Adaptive Network Path Restructuring Technique for Internet of Things (IoT)



Sharma Shamneesh, Manuja Manoj and Kishore Keshav

Abstract In the field of Internet of Things, self-adaptation and restructured network management have been a challenge since the inception of this field. Automatic node adjustments in the sensor fields are one of the key challenges in IoT. Gateway or Sink failure problem is another issue where researchers have to point their focus. In the present paper, we are providing a scenario where Self-Adaptive Node and Network Path Restructuring can resolve the issue of automatic restructuring of sensor network and redirect the data in case of node failure through automatic sharing of Routing State Table with the nearest neighbor gateway or sink. This solution will also provide an automatic service reconfiguration technique in case of service updates to be done on the entire network in a single step. This paper presents a technique of Network Path Restructuring to solve the node failure problem.

Keywords IoT · Network restructuring · Routing · Shortest path algorithm

1 Introduction

The IoT environment is different from the WSN environment [1] in terms of sensory data capturing, unique identification, and communication capabilities. In the recent technologies, IoT is using the gateway control centers, which manage the transfer

S. Shamneesh (✉)

Department of Computer Science and Engineering, Chitkara University, Chandigarh, Punjab, India

e-mail: shamneesh.sharma@gmail.com

M. Manoj

Vertical Head (IT) Inurture Education Solutions, Bangalore, India

e-mail: manoj_manoja@yahoo.in

K. Keshav

Department of Computer Science and Engineering, Alakh Prakash Goyal Shimla University, Shimla, Himachal Pradesh, India

e-mail: mails4keshav@gmail.com

of application data between gateways. During a situation of failure, IoT gateways can transfer their applications and device connections to another gateway at the time of a defined threshold limit. Now, the problem with this is that what will happen to the gateway if it fails without reaching the threshold value. To solve this problem, we propose a continuous backup of the state table along with data to the nearest neighboring gateway. So after the failure of gateway, the whole process will automatically be diverted to the defined nearest neighboring gateway. Researchers have started the inculcation of Artificial Intelligence in the Internet of Things (IoT). There are so many issues in IoT, which can be solved by using machine learning [2]. The basic idea behind the use of machine learning in the IoT is to reduce the unplanned downtime, increase the operational efficiency and enhance the risk management. The intersection of ML can be done at three levels while working with IoTs, i.e., Node Level (Hardware Perspective), Software Level (Operating System Perspective), and Network Level (Routing Algorithm Perspective).

1.1 Node Level

Node is generally fixed on the things or it is a thing itself. The restructuring of the IoT nodes (Sensor) is the concept of making the sensor node smarter at hardware level so that it can minimize its power consumption, disseminate the data on the most effective route, self-reconfiguration of primary and secondary data server or gateways in case of power failure and self-adaptation sensing capabilities during environment change [3]. The self-driving Car of Tesla is the individual unit that can be considered as a thing. Now, this car is adapting the knowledge from the environment and learning new routes and challenges which occur during the automated driving process.

1.2 Software Level

In the software level, we can automate the provisioning, maintenance and restructuring individual IoT component at OS and Software level. Every now and then, we come up with OS/Software updates. As an IoT engineer, we perform these tasks manually. Using the software level automation, we can easily perform these tasks without any human intervention.

1.3 Network Level

It is always mandatory that a thing in the IoT will be on a network and share the information with each other. Now to find out the best possible routes and restructuring of new best routes is the work that can be done on the Network Level. Today, the

IoT depends upon a centralized, client–server archetype to authenticate, authorize the multilevel users, and connect different nodes in a network. The current epitome is adequate for present IoT ecosystems where few things (devices) are implicated, but in the scenario where the networks will grow and fasten upon the hundreds of billions of devices, centralized model will twirl into a bottleneck.

2 Literature Survey

Machine Learning is a type of Artificial Intelligence, which endows machines with the capability to gain knowledge and extend it without any programming paradigm. In the paper [4], the authors have presented a taxonomy of machine learning algorithms, which explains that different techniques are functional to the data in order to extract a higher level of information. To make the IoT environment totally automatic, we need to make the nodes so smart that they can update themselves time to time. In the paper [4], the authors have applied Machine Learning techniques to the diverse networking contexts, which include congestion control, resource allocation, and network bottleneck detection. The techniques of [5] can be applied in the IoT sensor fields so that whenever, there will be any additional replacement of node in the network, it can update itself by using automatic resource allocation techniques. After the addition of the nodes, it is very important to check the probable delivery of the data to the gateway. In [6], the authors have proposed a scheme of application of machine learning (ML) algorithms to determine the probability of successful deliveries of packets from source to destination. As at the IoT sensor field, addition and deletion of node can raise a problem of vulnerability to Sybil attacks. In the paper, survey on Sybil attack defense mechanisms in wireless ad hoc networks [7], the authors discussed the scenario where malicious nodes can take the identity of any node in the wireless communication and take all the controls over data and communication. Andrea Zanella and Lorenzo Vangelista discussed in [8] by facilitating the undemanding admittance and communication with an extensive diversity of devices, the IoT will foster the development of a number of applications. This will further lead to make the implementation of the mammoth amount and accurate assortment of data generated by these devices to endow with new services. As suggested by the authors of [9], the development of general IoT architecture is a very intricate assignment because of the extremely large domain of devices, technologies, and services. Data handling is also one of the hardest tasks in large IoT network, so there is a need for smart data handling techniques in these networks. In [10], the authors discussed a technique of Accommodative Bloom Filter along with all its algorithms. A lot of work has been done on the routing algorithm implementation. The authors of [11] discussed various routing techniques for IoT along with implementations. A large IoT environment involves heavy data transmission causing bottleneck problems. Load balancing is an important technique in the routing, which can handle multiple users in the IoT environment. The author of [12] developed a technique of load-balanced routing that solved the problem of bottleneck in the large IoT net-

works. The authors of [13], proposed an epitome of opportunistic Internet of Things which encourages the opportunistic exploitation of communications between IoT devices in a network to attain the enhanced reliability, connectivity, competence, and network existence. Proper selection of the route in the networking is as important as the human brain functioning for the proper communication channel. Use of Machine learning in IoT has given a fair chance to the researchers to make the network more sustainable. Proper routing spans the network life, which is again an important issue in the IoT and WSN Environments. The authors of [14] have discussed the importance of machine learning in link cost in the routing of WSN. They have also performed the critical analysis of five machine learning algorithms for estimation of link cost. Various technologies of WSN have been used in the IoT so we can conclude a thing that WSN has been a framework for IoTs. This evolution has been discussed in the paper [15]. Along with routing, energy efficiency in WSN and IoT environments has been a remarkable research area where researchers have focused. In the paper [16], the authors had discussed the energy-efficient protocol in the WSN. Nowadays, machine learning is not the only technology where IoT is merged but it has also been merged with cloud, blockchain, cryptocurrency, and many more. In the paper [17], an algorithm has been proposed for IoT data dissemination using cloud services. As we all know that machine learning is the course of action of abolition of human involvements in the system. It is all about making the whole system working automatic [18].

Findings from Literature (Research Gaps):

- There are papers which are stating that there is a problem of node failure in the ad hoc networks which further lead to data loss.
- Some of the papers also state that the failures acknowledged on time can lead to the commencement of the recovery practices.
- The connectivity is the foundation of any IoT application or environment. Connectivity failure may lead to data or information loss. Missing data or information is the general occurrence in many IoT projects in the corporate systems.
- There are numerous numbers of implementations of IoT Systems. So, several architectures and systems are available across the enterprise.
- Normally, the reconfiguration of the nodes/things is done manually by the system administrator.

This paper is organized into six segments. Segment I is dealing with the introduction and Segment II is dealing with all the related literature, whereas Segment III states the methodology used. Segment IV is exclusively kept for the discussion of results. Segment V is focusing on conclusion of the research and future scope. This paper solved the problem of node replacement after node failure. After the replacement, the gateway will reconfigure the node with predefined configuration (user dependent) and automatic network adjustment.

3 Methodology

The smart IoT systems will automate the whole process for configuration and backup gateway selection in case of failure or maintenance by the virtue of which there will be a reduction in the planned or unplanned downtime. In this paper, we have simulated the IoT environment in a way that when a mobile node comes near to the transmission range of gateway Sink_2 or Sink_3, gateway initiates the *auto-reconfig SCRIPT* and the mobile node gets configured to the IoT Sensor field. In Fig. 1, the flowchart for the self-adaptive network path restructuring technique is discussed. At the initial level, the script will check whether the node is transmitting data or not. In the success condition, route information of the node is stored in the state table of the gateway node, and then the current information of the state table is replaced with the previous state table. In this way, the table becomes the current state table of the gateway node. In case of failure of data transmission at node level, the condition of node working

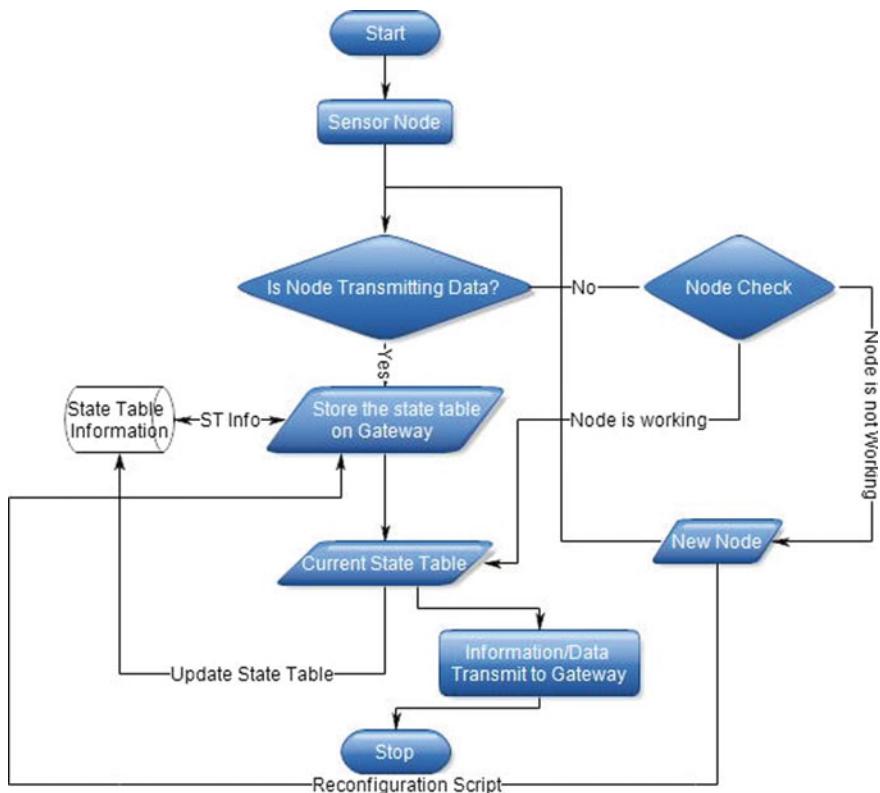


Fig. 1 Algorithm for node reconfiguration and path restructuring

is checked using a monitor script. If node is working, then the current state table is updated at node, otherwise new mobile node is inserted and its route information is updated in the current state table.

4 Experimental Setup

The whole simulation has been done on CupCarbon simulator. Although the simulator is platform independent, so we have used it on Ubuntu Linux [17.04] and JDK [1.8.0]. This simulator allows us to design, visualize, debug, and validate distributed IoT algorithms. It creates environmental scenarios such as IoT Things, and generally within educational and scientific projects. It is a platform-independent simulator written in Java.

In this paper, we have simulated the IoT environment for eight sensors devices (things) S1, S2, S3, S4, S5, S6, S7, and S8, which are transmitting data over Zigbee network protocol. There are two gateways Sink_2 and Sink_3, which are receiving data from these sensors. The sensors S1, S3, S5, and S7 are transmitting data to gateway Sink_2 and S2, S4, S6, and S8 are transmitting data to gateway Sink_3. For each component, the CupCarbon is generating the senscript code in (.CSC) file. A group of senscripts file represents all the components. To simulate the environment, out of eight devices in the field, one will get down due to unavoidable circumstances. In the present model, S6 will get down. Now to replace S6, we have inserted the mobile node M1 in the field. The mobile node took the charge of S6 and start transmitting data to Sink_3.

Figure 2 has defined the discussed scenario, in which S6 gets down and M1 starts working at the place of S6. Basically we can say that in real testbeds when one device gets down, we replace it with the newer or temporary node. Now after adding a new device in the field, we proposed a technique in which the added device will configure itself by asking for the configuration file from the gateway node and running it on its own OS and services. To achieve the second objective of load balancing of gateways, Sink_2 and Sink_3, Sink_3 gets down and all the nodes (S2, S4, M1, and S8) connected to Sink_3 start transmitting data to the Sink_2. The node is reconfigured using auto-reconfig script, which is invoked automatically upon a node failure or when a new node is introduced in the network. It assigns the state table values and a shortest possible path to the new node to the sink (Fig. 3).

5 Results and Discussions

The present paper states a scenario in which all solutions are present in case of the failure of nodes at field level. The Self-Adaptive Network Path Restructuring Technique had enabled the M1 to take immediate charge of S6 and started performing in the same way as S6 was performing earlier in the network. The following state

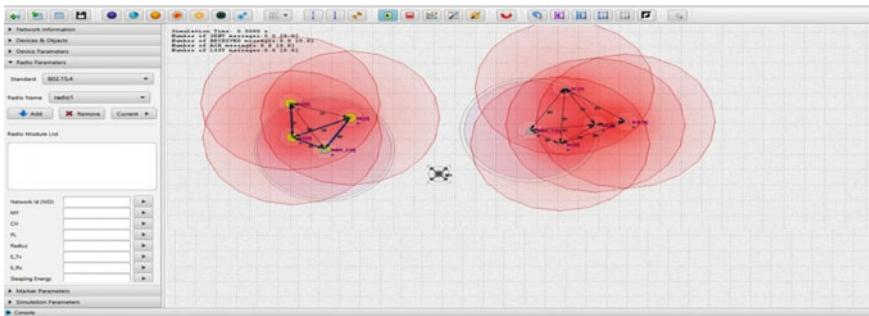


Fig. 2 Cupcarbon IoT simulation environment

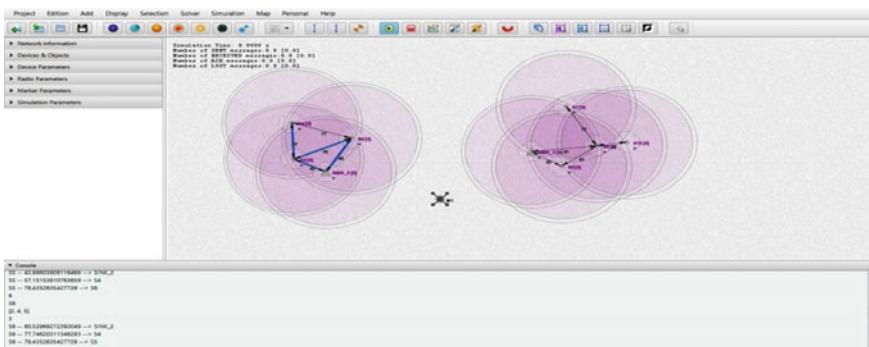


Fig. 3 IoT field where M1 acting as S6

table represents the performances of the gateway Sink_3 when S6 gets down. In the following graph, we have shown the state table and network performance of Sink_2 and Sink_3 in the normal conditions of the network. Each set in (3, 10, 7, and 9) represents the Sink and associated nodes with shortest possible path for data transmission. Sink_2 and Sink_3 can be synchronized using time-based *cron script*. Figure 4 shows the performance of the same gateway when M1 has been added at the place of S6 (Fig. 5).

6 Conclusion and Future Scope

The adaptive network and fail-proof connectivity are playing an important role in the development of robust IoT networks. This research work proved that applying Node-Level Network Path Restructuring Technique at both levels node as well as network increases the overall reliability of the network. Load balancing with automation can do miracles in the field of routing and it can manage the data loss in the IoT envi-

Fig. 4 Log of state table for Sink_2

```

▼ Console
S5 -- 43.98603808116469 --> SINK_2
S5 -- 57.15153810763659 --> S4
S5 -- 78.4352635427729 --> S6
6
S6
[2, 4, 5]
3
S6 -- 60.52969272392049 --> SINK_2
S6 -- 77.74620311346293 --> S4
S6 -- 78.4352635427729 --> S5
7

```

Fig. 5 Log of state table for Sink_3

```

▼ Console
S8
[3, 10, 7, 9]
4
S8 -- 81.36371098499222 --> SINK_3
S8 -- 37.81105928368628 --> S10
S8 -- 72.63504488209507 --> S7
S8 -- 53.47308878478493 --> S9
9
S9
[3, 8]

```

ronments. The replacement of the node again leads to a scope where we can enlarge the network by adding more nodes in a dynamic way. Reducing the human intervention (system administration) at node level provides an alternate approach toward fail-proof connectivity in the IoT environments. The present work only focuses on the performance of the network by adding new devices (nodes or gateways) in the IoT field. By machine learning at node level, we have achieved only the enhancement of the network but totally forget the security part where intruder nodes can also become the part of the network. Further research can be done on a mechanism where authentication-based method can be implemented while adding new nodes to the network. Research can be done on gateway-level failure and development of automated monitoring protocol for IoT management.

References

1. Kocakulak, M., Butun, I.: An overview of wireless sensor networks towards internet of things. In: 2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC), 02 March 2017. IEEE Press, USA (2017)
2. Sheth, A.: Internet of things to smart IoT through semantic, cognitive, and perceptual computing. *IEEE Intell. Syst.* **31**(2), 108–112 (2016). IEEE Press, New York
3. Sánchez-Sinencio, E.: Smart nodes of Internet of Things (IoT): a hardware perspective view & implementation. In: GLSVLSI ‘14, Proceedings of the 24th edition of the great lakes symposium on VLSI. ACM, New York (2014)
4. Mahdavinejad, M.S., Rezvan, M., Barekatian, M., Adibi, P., Barnaghi, P., Sheth, A.P.: Machine learning for internet of things data analysis: a survey. *Digit. Commun. Netw.* 161–175 (2018)
5. Valadarsky, A., Schapira, M., Shahaf, D., Tamar, A.: A machine learning approach to routing. *Netw. Internet Architecture, Comput. Sci. Sect.* (2017). Cornell University Library
6. Sharma, D.K., Dhurandher, S.K., Srivastava, R.K., Mohananey, A., Rodrigues, J.J.P.C.: A machine learning-based protocol for efficient routing in opportunistic networks. *IEEE Syst. J.* (2016)
7. Vasudeva, A., Sood, M.: Survey on Sybil attack defense mechanisms in wireless ad hoc networks. *J. Netw. Comput. Appl.* **120** (2018)
8. Zanella, A., Vangelista, L.: Internet of Things for smart cities. *IEEE Internet Things J.* **1** (2014)
9. Zhu, J., Song, Y., Jiang, D., Houbing Song.: A new deep-Q-learning-based transmission scheduling mechanism for the cognitive Internet of Things. *IEEE Internet Things J.* 2327–4662 (2017)
10. Singh, A., Garg, S., Batra, S., Kumar, N., Rodrigues, J.J.P.C.: Bloom filter based optimization scheme for massive data handling in IoT environment. *Future Gener. Comput. Syst.* (2017)
11. Triantafyllou, A., Sarigiannidis, P., Lagkas, T.D.: Network protocols, schemes, and mechanisms for Internet of Things (IoT): features, open challenges and trends. *Wirel. Commun. Mob. Comput.* **2018** (2018). Article ID 5349894
12. Tseng, C.H.: Multipath load balancing routing for Internet of Things. *J. Sens.* **2016** (2016). Article ID 4250746
13. Kim, T., Lim, J., Son, H., Shin, B., Lee, D., Hyun, S.J.: A multi dimensional smart community discovery scheme for IoT-enriched smart homes. *ACM Trans. Internet Technol.* **18**(1) (2017). Article 3
14. Singh, K., Kaur, J.: Machine learning based link cost estimation for routing optimization in wireless sensor networks. *Adv. Wirel. Mob. Commun.* **10**, 39–49 (2017). ISSN 0973-6972
15. Kishore, K., Sharma, S.: Evolution of wireless sensor networks as the framework of Internet of Things-a review. *Int. J. Emerg. Res. Manag. Technol.* **5**(12) (2016). ISSN: 2278–9359
16. Sharma, S., Mathur, R.P., Kumar, D.: Enhanced reliable distributed energy efficient protocol for WSN. *IEEE Int. Conf. Commun. Syst. Netw. Technol.* (2011). IEEE Press, New York
17. Sharma, S., Kishore, K.: Data dissemination algorithm using cloud services: a proposed integrated architecture using IoT. In: 2nd International Conference on Innovative Research in Engineering Science and Technology (IREST-2017), Eternal University, Baru Sahib, Sirmour (H.P.), India, 7–8 April 2017 (2017)
18. Liu, X., Zhang, X., Guizani, N., Lu, J., Zhu, Q., Du, X.: TLTD: a testing framework for learning-based IoT traffic detection systems. *Sensors* **18**, 2630 (2018)
19. Dhurandher, S.K., Borah, S.J., Woungang I., Bansala, A., Gupta, A.: A location prediction-based routing scheme for opportunistic networks in an IoT scenario. *J. Parallel Distrib. Comput.* (2017)

Design of Pole Placement-Based State Feedback Controller for a Impulse Hydro Turbine



Shubham Ashish, Sagar Saraswat, Sagar Garg and Alok Kumar Pandey

Abstract The paper is presenting an approach to control the turbine dynamics in order to control the electricity generation. Our challenge is to control the discharge or the flow rate with the use of electrical controllers. A pole placement (PP) based state feedback controller (SFC) has been proposed in this paper to stabilize the system in MATLAB. Traditionally, the flow is controlled with the help of governor as a mechanical controller. Here, turbine dynamics include the height of dam, value of gate and the angular speed of the rotating shaft of turbine, which is coupled with the rotor of alternator.

Keywords Impulse hydro turbine · PP technique · SFC

1 Introduction

The water of flowing river is obstructed with the help of a structure called dam. The dam basically stores water up to a certain height in the form of potential energy. The water through the dam is allowed to flow through a pipe called penstock on the turbine blades; here, the potential energy changes into the kinetic energy. The shaft of rotating turbine is coupled with the synchronous generator. The extra wasted water is exited with the help of gate and returned to the reservoir [1].

Conversion of energy can be done from mechanical to electrical with turbine generator combination as shown in Fig. 1.

h = Pressure, T_m = Mechanical Torque, T_e = Electrical Torque, G = Gate,
 q = Flow rate.

There is much irregularity on the load side due to variations in day and night or seasonal consumption. These variations in the load side cause different amounts of

S. Ashish (✉) · S. Saraswat · S. Garg · A. K. Pandey
Electrical and Electronics Department, KIET, Ghaziabad, India
e-mail: shubham.1521149@kiet.edu

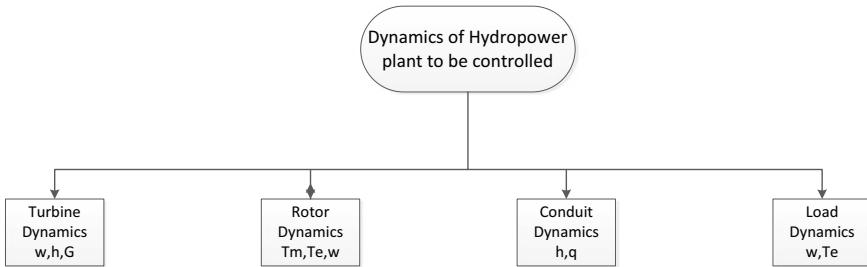


Fig. 1 Dynamics of hydro turbine to be controlled [2]

electricity to be generated putting pressure on the load side. The variation in load causes variation in the water discharge through the penstock to generate different amounts of electricity.

This variation put more emphasis on the turbine dynamics as well as on the need of controllers. Like PID or methods like pole placement in which poles can be adjusted to control the system or make it more stable [2]. For hydropower generation, turbines can be selected optimally like power generating modules are used to maximize energy. Hydropower plant models test of hydropower plant to detect grid fault-4 unit at 112 KW with $2 * 330$ kV [3]. Three-phase short circuit to ground in high voltage side.

Hydro turbine governor controlling methods: An analysis has been done to control the turbine governors and grading in existing electrical power system. It shows the processes involved, tuning and their limitations [4].

Design of speed control system for Pelton turbine: Quantity of water flowing, nozzle and deflector are the main parts of regulating mechanism and speed control system. Designing of Pelton turbine and varying the jet diameter, nozzle velocity and deflection is done in order to vary the governor speeds [5].

Energy generated from the head drop of hydropower has 1000 times weight than wind power. A reaction turbine spins faster when its head is constant and the diameter of rotor affects the maximum speed of rotor [5]. A review on Advances in Design and analysis of draft tube for reaction turbine, draft tubes are conduits of gradually increasing cross sectional area. Pressure head is recovered as water flows down draft tube. It is observed that CFD can be used to predict performance of draft tube and save time, cost and resources [6]. Enhancement in power generation of hydropower plant: Discharge of water is minimum when a turbine operates at maximum efficiency. The load of unit in power plant depends on the load demand while the head on turbine depends on water level in reservoir. Numerical computation on constant load is done to minimize discharge and constant head to maximize generation. Conditional maintenance (Naval Science and Technological Laboratory): Use of PMP as a maintenance strategy in power industry. Through PMP, it has been possible to reduce 7% reduction in transmission and distribution losses [7].

2 System Description

The proposed system is of hydropower plant is installed at the generating station. There is a discharge of water through dam by a pipe called penstock striking on the turbine causing it to rotate, this results in the conversion of one form of energy into another form of energy. The input to the system includes the discharge of water or the flow rate and on the output side, we are controlling the angular speed of the shaft of the turbine. The potential energy stored in the dam varies with the height of the dam. The flow rate can also be controlled with the help of the gate. Through the given system, we can generate the electricity in correspondence to the variation in the load. Control of plant has four major components: 1. Penstock, 2. State feedback controller, 3. Turbine and 4. Generator. Here, we use pole placement-based state feedback controller, which reduces the settling time (t_s) to reduce the Maximum overshoot (M_p) to get a stable step response.

The assumptions for the given hydropower plant system are the following:

1. The disturbance to the system is assumed to be zero. Hence, the matrix D will be zero.
2. We are using a linearized system model for analysis Table 1.

The state-space representation of the turbine dynamics for hydro power plant is given below [2]

$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$y(t) = Cx(t)$$

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & \frac{-1}{t_p} & 0 & 0 \\ 0 & \frac{1}{t_g} & -\frac{1}{t_g} & 0 \\ 0 & -\frac{2}{t_g} \left(\frac{2}{t_w} + \frac{2}{t_g} \right) & -\frac{2}{t_w} & 0 \\ 0 & 0 & 0 & \frac{1}{m} \end{bmatrix}, B = \begin{bmatrix} 0 \\ -\frac{1}{t_p} \\ 0 \\ 0 \\ 0 \end{bmatrix}, C = [1 \ 0 \ 0 \ 0 \ 0]$$

Table 1 Details of parameters [2]

Symbol	Name	Values (s)
M	Governor time constant	6
R_p	Permanent droop	0.5
T_w	Water start up time	1.3
T_g	Gate servo time constant	0.2
T_p	Pilot valve time constant	0.02
d	Droop time constant	1

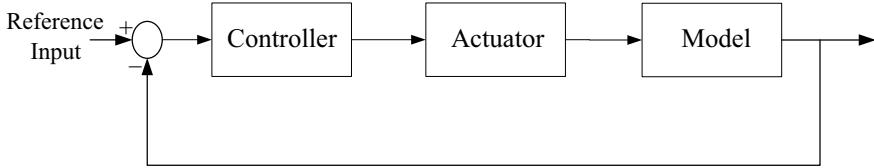


Fig. 2 Basic control system configuration [9]

3 Controller Design Method

3.1 Pole Placement Method

Turbine dynamics have five open loops $\{0, -49.22, -0.3988+0.6943i, -0.3988-0.6943i, -5.9846\}$. Out of all two poles of turbine are on origin, so the system is critically stable. Pole placement-based state feedback technique is used to stabilize the system. Poles have been chosen according to design specification so new location of poles are $\{-4.5+4.5i, -4.5-4.5i, -50, -1, -6\}$. The necessary and sufficient condition for the placement of closed-loop poles in the complex plane is that the system is controllable, If all n state variables x_1, x_2, \dots, x_n can be accurately measured at all times, according to linear control law of the form [8]

$$u(t) = -k_1x_1(t) - k_2x_2(t) - \cdots - k_nx_n(t) = -kx(t) \quad (2)$$

where K_1, K_2 are state feedback gain matrix as shown in Fig. 2.

4 Simulation Result

One of the poles of the system is on origin so the system performance is not stable enough that is shown in Fig. 3.

Response of system with pole placement-based state feedback controller is shown in Fig. 4. It is showing a stable response.

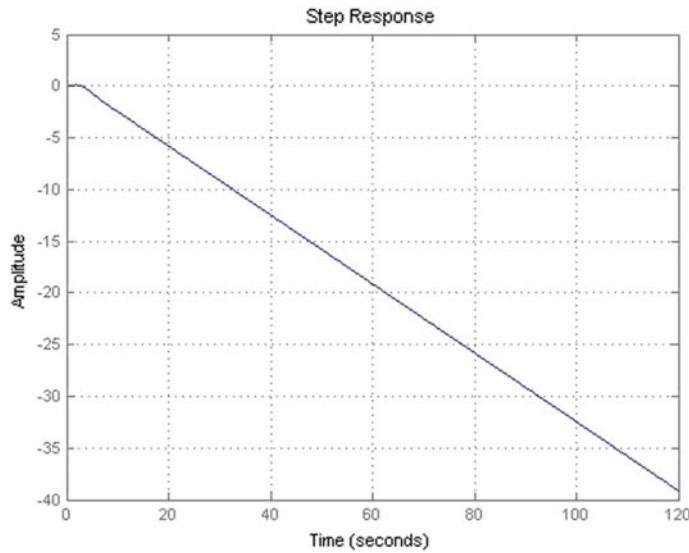


Fig. 3 Response of turbine without controller

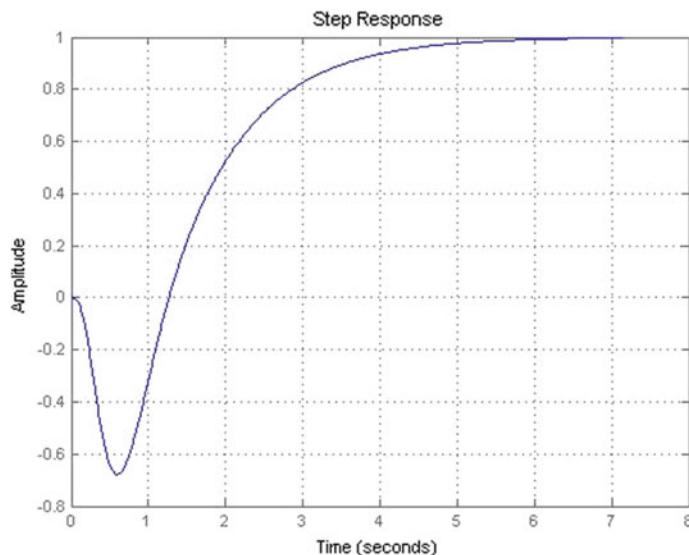


Fig. 4 Response of turbine with controller

5 Conclusion

Pole placement-based state feedback controller has been designed to stabilize the turbine dynamics, which was not stabilized because of its one pole on origin. Poles of system have been placed in such a way that response of system is able to match the design specifications. LQR-based controller can be done in future to get the optimum results.

References

1. Sangal, S., Garg, A., Kumar, D.: Review of optimal selection of turbines for hydroelectric projects. *Int. J. Emerg. Technol. Adv. Eng.* (2013)
2. Ding, X., Sinha, A.: Sliding mode/H1 control of a hydro-power plant. In: American Control Conference on O'Farrell Street, San Francisco, CA, USA (2011)
3. Elbatran, A.H., Yaakob, O.B., Ahmed, Y.M., Shabara, H.M.: Operation, performance and economic analysis of low head micro-hydropower turbines for rural and remote areas: a review. *Renew. Sustain. Energy Rev.* (2015)
4. Culberg, J., Negnevitsky, M., Kashem, K.A.: Hydro-turbine governor control: theory, techniques and limitations. In: Australasian Universities Power Engineering Conference (AUPEC 2006) (2006)
5. Jade, N.P., Raut, S.S., Wadnere, G.A.: Performance analysis of parameters related to low head hydro turbines & their effect on efficiency of turbines. *Int. J. Eng. Technol. Manag. Appl. Sci.* **3**(4), ISSN 2349-4476 (2015)
6. Rai, V., Saxena, N.V.: A review on advances in the design and analysis of draft tube for reaction turbines. *Int. Res. J. Eng. Technol. (IRJET)* **05**(05) (2018). e-ISSN: 2395-0056, p-ISSN: 2395-0072
7. Rao, V.B.: SC 'G' Condition Based Maintenance for improving Power Plant Economics. Naval Science & Technological Laboratory
8. Pandey, A.K., Shukla, A., Gupta, A., Gangwar, M.S.: Linear flight control of unmanned aerial vehicle. In: Proceeding of ICICCD 2017, Advances in Intelligent Systems and Computing. Springer ASIC Book Series, vol. 624, pp. 924–931 (2018)
9. Rawat, D., Bansal, K., Pandey, A.K.: LQR and PID design technique for an electric furnace temperature control system. In: Proceeding of International Conference on Intelligent Communication, Control and Devices Part of Advances in Intelligent Systems and Computing. Springer ASIC Book Series, vol. 479 (2016)
10. Pandey, A.K., Chaudhary, T., Mishra, S.: Longitudinal control of small unmanned aerial vehicle by PID controller. In: Proceeding of ICICCD 2017, Advances in Intelligent Systems and Computing. Springer ASIC Book Series, vol. 624, pp. 393–400 (2018)

Paradigm-Based Morphological Analyzer for the Gujarati Language



Dikshan N. Shah and Harshad Bhadka

Abstract A morphological analyzer is a tool, which makes a syntactic analysis of a word and obtains root form of an inflected word form. The elementary step to identify a given sentence is the morphological analysis. We portray Morph analyzer for the Gujarati language. For a superior understanding of a language, word level, sentence level, context level and discourse level analysis has to be done. One of the tasks is the morphological level analyzing for various word forms. In this paper, we have discussed a paradigm-based approach for morphological analysis for various POS tags. The algorithmic development gives an accuracy of analysis for noun 84.50%, 81.50% for verb and 80.50% for adjectives.

Keywords Morphology · Paradigm-based approach · Inflectional form · Morphological analysis

1 Introduction

Natural language generation systems translate information from computer databases into human languages. The fundamental goal of Morphological analyzer is to clarify a system of language words, relations and conceptual information, which is used by computer logic to execute artificial language interpretation.

D. N. Shah (✉)

Gujarat Technological University, Ahmedabad, Gujarat, India

e-mail: dikshan817@gmail.com

H. Bhadka

Faculty of Computer Science, C U Shah University, Wadhwan, Gujarat, India

e-mail: harshad.bhadka@yahoo.com

2 Existing Work

Analysis of Hindi nominal inflection word showed using Distributed Morphology. In their [1] paper, they mentioned Hindi Morphology, inflection and its execution in the Distributed Morphological Analyzer. They [2] have done a morphological analysis of various parts of speech such as noun, pronoun and adjective for NLization of Punjabi language with EUGENE tool. They described three categories of morphology: (1) Attribute label resolution morphology, (2) Relation label resolution morphology and (3) Noun, Adjective, Pronoun morphology. For the corpus of 500 sentences, they trained 100 sentences and 90 sentences were correctly identified and they achieved 0.90 as F-measure value. By using Stemmer, they [3] identified inflection in appeared in most Gujarati text. EMILLE corpus was used for stemmer's performance and achieved 92.41% accuracy. Morphological rules were developed for Gujarati language classes and lexicon databases. They [4] developed lexicon-based Unicode system with more than 15,000 words. They [5] mainly work on segmenting inflected words into its root word and its connected morphological elements along with features stipulating the morphological structure. A principal subject of examination [6] has been carried out for the verb inflectional system in the Colloquial Bangla (CB) language, which is considered distinctive from the Standard Bangla (SB) with interest to the forms of the grammatical markers. A model has been presented for Kannada inflectional morphological analysis. The model was designed [7] using rule-based, paradigm-based, questionnaire-based and affix-stripping approaches. Experimental results were achieved with 90% accuracy. They [8] have shown rules based on the morphological analyzer. By taking Gujarati sentences as an input, it generates grammar classes, number, gender, tense and person information with its root words. They have taken input from essays and short stories as input and obtained 87.48% accuracy upon evaluation.

3 Gujarati Morphology

The language smallest grammatical unit is called a morpheme. Gujarati is a morphologically rich language [1]. Based upon the similarity in grammatical features and word formation process, words are categorized into various paradigms.

3.1 *Morphology in Linguistics*

Morphology describes the internal structure of words. While words are generally accepted as being the smallest units of syntax. In most languages (not in all), words are related to other words by rules. For example, છોકરા (Chōkarā)— છોકરો (Chōkaro)— છોકરું (brat) is relatively closed.

In Gujarati, language narrators identify these relations from their inferred knowledge of the rules of word formation. Even they perceived: નગર (Nagara)—town and નગરો (Nagarō)—towns are same as પત્ર (patra)—letter and પત્રો (Patrō)—Letters. Same as નગર (Nagara)— નગરવાસી (Nagaravāsī), ભારત (Bhārata)— ભારતીય (Bhāratīya).

3.1.1 Gujarati Lexemes and Word Forms

The peculiarity between the two senses of the word is possibly the most significant one in morphology. In the following example: નગર (Nagara)—town and નગરો (Nagarō)—towns are the same word called ‘Lexeme’. Similarly, in the second example: નગરવાસી (Nagaravāsī), ભારત (Bhārata)— ભારતીય (Bhāratīya) refers to two different entities called word forms.

3.1.2 Inflectional Versus Derivational Morphology

Inflectional morphology is the study of those processes of the word formation where new words with different forms but the same meaning are formed from an existing stem [2]. For example, Plurals: બસ (Basa)—Bus—બસો (Basō)—Buses, કાર (Kāra)—Car— કારો (Kārō)—Cars past tense: do (કર્યું) —> did (કર્યું), Sing (ગાયું) —> Sang (ગાયું).

Aspect: child (બાળક) —> childhood (બાળપણ). Derivational morphology is the learning of the processes of the word formation where new words are formed from the existing stems through the addition of morphemes. A very common type of derivation in Gujarati is the formation of new nouns, often from verbs or adjectives or the nouns. It is called nominalization [7]. For example, નસીબ (Nasība)—Luck—> નસીબદાર (Nasībadāra)—Lucky, મહાન (Mahāna)—Great—> મહાનતા (Mahānatā)—Greatness.

3.2 Paradigms and Morph Syntax

The comprehensive collection of connected word forms related with a dedicated lexeme is called Paradigms. The word forms of a lexeme are organized as per inflectional types such as number, gender, tense, case, mood and aspect. For example—personal pronouns are ordered into categories of number (singular & plural), the person (first, second and third), case (subjective, objective and possessive) and gender (masculine, feminine, and neuter).

Table 1 Noun morphemes

Id	Root Word	Suffix	Example
1	કુતરુ (Kutru)	ગું, અંલ, અંગ, અંન, અંણો, અંણો	કુત્રાએઓ (Kutrao)
2	કટકો (Katko)	અંલ, અંલ, અંગ, અંન	કટકો (Katko)
3	વાક્ય (Vaakya)	અંલ	વાક્યો (Vaakyo)

[Table – 3.1 Noun Morphemes]

3.3 Gujarati Word Classes

The first step in developing a morphological analyzer is to describe the word classes and grammatical information, which will be required for words of these word classes. It is not possible for all nouns to follow a same inflectional pattern. So, it is more important to find various paradigms for those word classes and group them according to those paradigms. Various word classes are shown herewith.

3.3.1 Noun Morphology

Nouns in Gujarati are transformed into gender, number and case. The Gujarati language has three genders: (1) masculine gender, (2) feminine gender and (3) neuter gender and two number types—singular and plural and three cases: (1) nominative, (2) vocative/oblique and (3) locative [3].

Morphological Rules for Noun:

- (1) If a word belongs to masculine gender and ends with ં (o), its singular form must terminate with ં (o) and plural form ends with ંાં-ં (aa -o).
For example, પંખો (Pankho) Fan— પંખોઓ (Pankhao) Fans.
- (2) If the word belongs to masculine gender but doesn't end with ં (o), its plural form must end with ં (o). For example, મશીન (Masin) Machine— મશીનો (Masino) Machines.
- (3) If the word belongs to feminine gender, its plural form must end with ં (o).
For example, સપાલી (Tapali) Postman સપાલોઓ—(Tapalio) Postmen.
- (4) If the word belongs to neuter gender and ends with ંાં (un), its singular form must end with ંાં (un) and plural terminates with ંાંંં (aaon).
For example, કપડ (kapadum) Cloth— કપડાં (kapadam) Clothing/ કપડાંઓ (kapadamo) Clothes.
- (5) If the word is in neuter gender but doesn't end with ંાં (un), its plural form must end with (o). For example, વેંગાણ (vengana) Brinjal— વેંગાણાં (venganaao) Brinjals.

In above Table 1, Root Word, and suffix list is used to detect this concept, and the final column shows one inflected word belonging to the same paradigm.

Table 2 Pronoun morphemes

English pronoun	Gujarati pronoun	Transliteration
I (First Person—Singular)	હું	huM
We (First Person—Plural)	આમે	(Ame)
We (First Person—Plural)	આપણે	(ApaNe)
You (Second Person—Singular)	તુ	(tuM)
You (Second Person—Plural)	તમે	(tame)
He/She/It (Third Person—Singular)/This(Masculine/Feminine/Neuter Gender)/These(Masculine/Feminine/Neuter Gender)	આ	(Aa)
He/She/It (Third Person—Singular)/That(Masculine/Feminine/Neuter)/Those (Masculine/Feminine/Neuter)	એ તે એ	(e) (more polite than (te))
They (for a person) (Third Person—Plural)	તેમો	(to)
They (for the non-living thing) (Third Person—Plural)	એ	(e)

Table 3 Some Adjective morphemes

Id	Root Word	Suffix	Example
1	હાલવું (Halavum)	ઓ, ઓની, ઓ	હાલવો (Halavo)
2	સુંદર (Sundar)	No Inflection	સુંદર (Sundar)
3	અભિમાન (Abhiman)	ઓની	અભિમાની (Abhmani)

[Table – 3.3 Some Adjective Morphemes]

3.3.2 Pronoun Morphology

Pronouns are used in place of a noun. Gujarati personal pronouns also distinguish three persons (first, second and third) in two numbers (singular and plural). Personal pronouns gain many cases, which includes nominative, ergative, accusative/dative, genitive and locative and instrumental (Table 2).

3.3.3 Adjective Morphemes

Adjectives are categorized into variant and non-variant adjectives grounded on the inflections which they conquer. A non-variant adjective does not inflect with gender (Table 3).

Morphological Rules for Adjective:

- (1) If the word belongs to neuter gender and ends with ઊં (uun), its adjective form of a word must terminate with ઊં (o), ઊં (ee), ઊં (u).
For example, હલ્કુ (Halku)— હલ્કો (Halko)— હલ્કી (Halaki)—Lightweight.
- (2) If the word is neuter gender and cannot be inflected then word itself consider as adjective form. For **example**, ખ્રાબ (Khrab)—Khrab (Khrab).
- (3) If the word is in neuter gender then its adjective form ends with ઊં (ee).
For example, ગમંડ (Gamand)— ગમંડી (Gamandi).

3.3.4 Verb Morphemes

Inflected forms of verbs in Gujarati are divided into three morphemes arranged in a definite sequence to form a morphological construction, namely: (1) the root morpheme which is associated with the basic meaning of the verb, (2) a verbal suffixes associated with the meaning of tenses and moods/aspect and (3) inflectional morphemes expressing the person number or the gender number of the subject [4].

Example: A verbal form like ‘કર્યું (karyu)’ - (did) is analysed as consisting of the root morpheme ‘કરી દુ (kari do)’. The suffix -yu- indicating the past tense and -u- indicating the singular number and neuter gender. Same way the suffix ‘-i- do’ indicating present tense and ‘(દુ) do’ indicating plural the number and neuter gender.

$$\text{Verb Morpheme equation} = R([c_1(c_2)]) = \left(\begin{bmatrix} Tf + No.p \\ \frac{Tp}{Tipr} + GNo \\ +No.p \\ TM + No.p \\ A \end{bmatrix} \right) \quad (1)$$

where R = Root, C₁ = Causative 1, C₂ = Causative 2, Tf = Future Tense, No. p = Person number, Tp = Past Tense, Tipr = Past Imperfect, TM = Tense and Mood, GNo. = Gender Number and A = Aspect.

Table 4 Verb morphemes

Id	Suffix	Root Word	Example
1	આવ્યું, વ્યું, આવશે, આવ્યા, એ, ઈશ, જો, ઓ, વું, તો, તી.	કરું (Karvu) ઉંઘ (Ungh)	કરશે (Karshe) ઉંઘતી (Unghati)
2	એ, ઈશ, જો, ઓ, વું, તો, તી, આડે, આડી, આડશે, અડયા	જમું (Jamu) ખા (kha)	જમાડું (Jamadyu) ખાઈશા (Khaisa)
3	અ, જો, એ, વું, તો, તી, ઇ, ધૂં, વ્યું, આવશે, આવ્યા,	ખાવું (Khadhu) કર (kar)	ખાસું (Khadhu) કરતો (karto)

[Table – 3.4 Verb Morphemes]

Possible verb morphological constructions are as follows:

1. R	કર (kar)	'do' (imperative)
2. R + C ₁ (imperative)	કરાવ (karav)	'cause to do'
3. R + C ₁ + C ₂ (imperative)	કરાવડાવ (karav-dav)	'make to cause to do'
4. R + Tf + No.p	કરીશુ (kar-is-u)	'We will do'
5. R + Tp + GNo. singular)	કર્યુ (karyu)	'did' (neuter gender,
6. R + Tipr + GNo.	કરતો (kar-to)	'he used to do'
7. R + No.p	કરુ (karu)	'I do'
8. R + TM + No.p future imperative)	કરજો (karjo)	'you do' (2 nd person,
9. R + A	કરત (karat)	'conditional'

Following Table 4, describes various Verb paradigms [4].

4 Morphological Analysis

The Gujarati language requires the verb in a sentence to develop an inflectional word form, which corresponds to a person and a number of the subject [8] (Table 5).

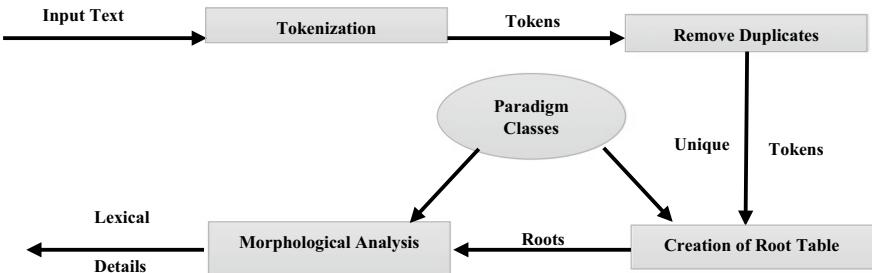
5 Methodology

To achieve the aim of designing, Paradigm-based approach is a simple and easy to use interface of the morphological analyzer for the Gujarati language (Fig. 1).

In this first module, convert inputted text sentences into word level tokens entailing of words without punctuation marks and other symbols called preprocessing of text. In this, it prepares the list of words and removes the duplicates. The second module

Table 5 Some verb morphemes

Case 1	Case 2	Case 3
Input word = મધમાયીએના	Input word = પત્રકારીએઓએ	Input word = લખશૈ
Case Marker = ના - મધમાયીએ	Case Marker = એ - પત્રકારીએ	Case Marker = NULL
Number Marker = એ - મધમાયી	Number Marker = એ - પત્રકારી	Number Marker = NULL
Fem. Gender marker = ઊ - મધમાય	Fem. Gender marker = ઊ - પત્રકાર	Feminine Gender marker = NULL
Suffix = ઊ - એના	Suffix = ઊ - એઓએ	Result1: Noun not found
Stem word= મધમાય	Stem = પત્રકાર	Verb Rule: ઊ શા - લખશુ
Result1: Category = NN Feminine Gender Plural	Result1: Category = NN Masculine Gender Plural	Stem = લખ
Result2: Verb not found	Result2: Verb not found	Result2: Category= VM.FUTURE Singular Gender

[Table – 4.1 Some Verb Morphemes]**Fig. 1** Architecture of the system

creates a table comprising of roots along with its classes which they follow. The third module analyses each word and provides morphological details of the input sentences.

5.1 Dictionary Generation

To create a language dictionary, a large Gujarati text file is considered as an input. Following is the algorithm steps for dictionary generation:

Algorithm—5.1 Steps to create dictionary

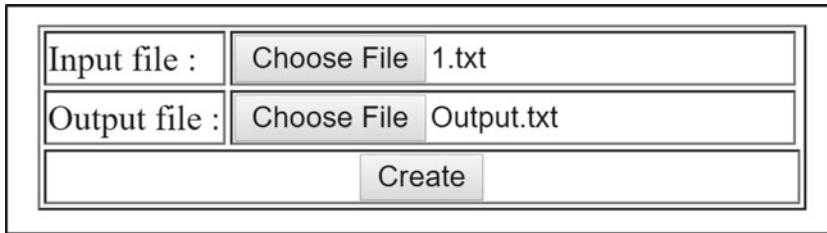


Fig. 2 Snapshot of Tool after selecting input file

મુખ્ય, તા. ૩૦ ઓક્ટોબર ૨૦૧૭ વિશ્વનું સૌથી ધનાઢ્ય કિકેટ બોર્ડ બીસીસીઆઇ ક્રિકેટરોને નાણા યુક્તવામાં 'કાપ' મૂકે છે? કમિટી ઓફ એડમિનિસ્ટ્રેટર્સ (સીઓએ)ને તપાસ કરતા અમ માલૂમ પડ્યું છે. નિયમ અનુસાર બીસીસીઆઇએ વાર્ષિક કોન્ટાક્ટ ધરાવતા વૈયરને પ્રસારણ અધિકારમાંથી થાથી આવકના ૨૬% રકમ આપવી જોઈએ. જેના સ્થાને બીસીસીઆઇ વૈયરોને માત્ર ૮% રકમ યુક્તવી રહ્યું છે. બીસીસીઆઇ આ રકમ ઐલાડીઓને બોનસ-પગાર તરીકે યૂક્વે છે.

Fig. 3 Input file (1.txt)

Algorithm

Step 1: Input Gujarati Text

Step 2: Tokenize Input text

Step 3: Append tokens to the output file

Step 4: Remove duplicate words

Step 5: Output words to file

According to algorithm steps [5], the input is passed as a file comprising of Gujarati strings. By applying tokenization on to the file, punctuation marks and other symbols will be removed. Generated tokens will be appended into the output file and duplicate tokens will be removed from the output file by creating a HashSet. This created HashSet is a collection of unique words which is known as a dictionary generation tool.

File 1.txt is an input file which comprises a collection of Gujarati sentences shown below and Output.txt is a file selected for transferring an output after cleaning the text (Fig. 2).

The following File 1.txt is an input file, which comprises of a collection of Gujarati sentences (Fig. 3).

By clicking of creating button, a text file will be created which contains all the Gujarati words in the input text file and cleaning the text by removing punctuation marks. Following output.txt file contains unique Gujarati words from 1.txt (Fig. 4).

5.2 Root Table Creation

These tables will be created by using a word list, we will get a list of roots along with.

મુંબિય	તા	૩૦	ઓક્ટોબર	સોમવાર	2017
વિશ્વનું	સૌથી	ધનાઢ્ય	કિકેટ	બોડ	બીસીસીઆઇ
કિકેટરોને નાણા	યુફ્લવામાં	'કાપ'	મૂઠ	થે	કમિટિ
ઓફ	એકમિનિસ્ટ્રેટર્સ	(સીઓએ)	તપાસ	કરતા	અમૃ
માટ્રામ	પડ્યું	નિયમ	અનુસાર	વાર્ષિક	કોન્ટ્રાક્ટ
ધરાવતા	ઘેયર	પ્રસારણ	અધિકાર	થતી	આવકના
૨૬%	રકમ	આપવી	જોઈએ	જેના	સ્થાને
માત્ર	૮%	ચૂકવી	રહ્યું	બીનસ-	પગાર
તરીકે	ચૂકવે				

Fig. 4 Output File (output.txt)**Table 6** Root Table

Root_name	Parameter_name	Root_name	Parameter_name
ગોપકી	નદી	પુસ્તક	ધર્મિયાળ
કપડાં	ટોપવાં	ઝડપ	રાત
રસ્તાં	ખંખા	વાર	કાર
બકરા	વાંદરા	સ્ત્રી	નદી

Algorithm—5.2 Algorithmic steps for creating root table

```

For each word W in word form list do
  For i ∈ { 0 ... W.length } do
    Take suffix S of W having length i
    For each paradigm class P do
      For every match of S with X; where X ∈ { four suffixes of P } do
        Set t := no. of characters to be removed from P before adding X to it to make a new word
        Set new_root_word := W - Suffix X + Suffix of P having length t
        For each Y of the four suffixes of P with U being the corresponding no. of characters to be removed, do
          Set new_word := new_root_word - Suffix of new_root_word of length + Suffix Y
          If new_word ∉ word form list
            Goto step 8
          End if
        End for
        Write entry in the root table in which new_root_word follows paradigm P
      End for
      End For
    End For
  
```

their paradigm name which they follow. Using this root and all other four suffixes of the class, four words are created. If all these words exist in the word form list, then the entry in the root table is made specifying the new root with its paradigm class (Table 6).

5.3 Morphological Analysis

The morphological analysis takes one word at a time and checks which paradigm it belongs to. It gives the lexical details as an output using the paradigm class.

Algorithm: 5.3 Algorithm for Morphological Analysis

```

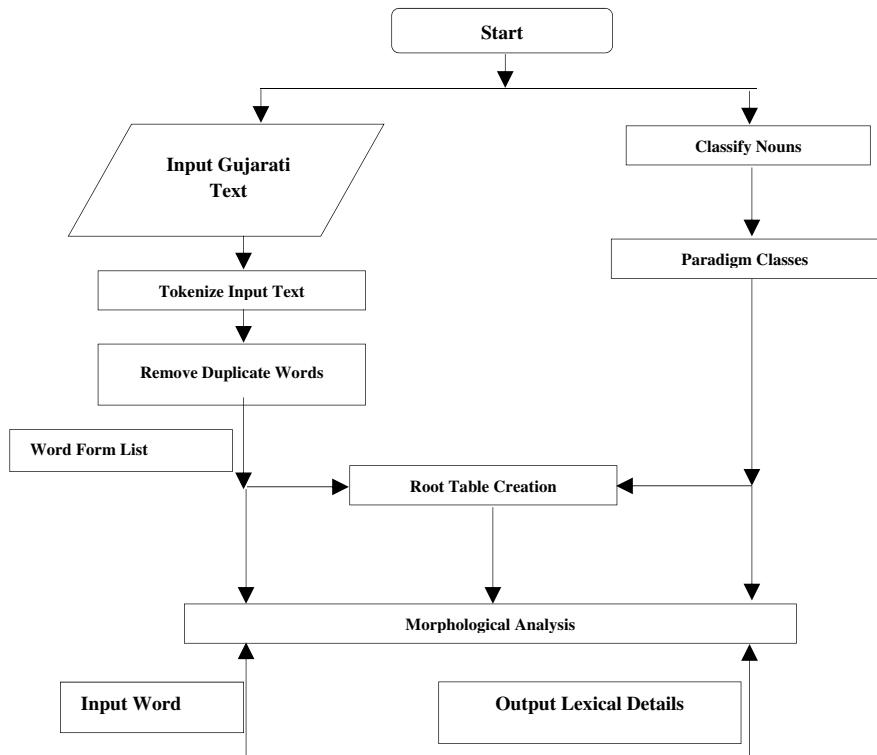
 $L := \text{empty set}$ 
 $\text{For } I = 0 \text{ to length of } W \text{ do}$ 
     $\quad \text{Let } S = \text{suffix of length } I \text{ in } W$ 
     $\quad \text{For each paradigms table } P$ 
         $\quad \quad \text{For each entry } B \text{ (consisting of a pair) in } P \text{ do}$ 
             $\quad \quad \quad \text{If } S = \text{suffix in entry } B \text{ then}$ 
                 $\quad \quad \quad \quad R = \text{root of paradigm table } P$ 
                 $\quad \quad \quad \quad J = \text{number of characters to be removed as shown in } B$ 
                 $\quad \quad \quad \quad \text{proposed\_root} = (W\text{-suffix } S) + \text{suffix of } R \text{ consisting of } J \text{ characters}$ 
                 $\quad \quad \quad \quad \text{If (proposed\_root is in DR) and (the root has paradigm } P)$ 
                     $\quad \quad \quad \quad \quad \text{Then construct a lexical entry } l \text{ by combining (a) features given in DR with the}$ 
                     $\quad \quad \quad \quad \quad \text{proposed\_root, and (B) features associated with } E.$ 
                     $\quad \quad \quad \quad \quad \text{Add } l \text{ to Set } L.$ 
                 $\quad \quad \quad \quad \text{End If}$ 
             $\quad \quad \quad \text{End If}$ 
             $\quad \quad \quad \text{End for every } B \text{ in } P$ 
             $\quad \quad \quad \text{End for every paradigm}$ 
         $\quad \text{End For}$ 

```

In the first step, take an empty set L in which we will enter all the morphological details of the input word W. Now, put the suffix of W in S of length 0, 1, 2, ... up to the length of W. Now, search S in every entry of B of the paradigm table P. When a match is found, fetch the root of that paradigm table as R. Find the proposed root by removing the suffix S from W and adding the suffix of R of same length as S. If this proposed root is found in the dictionary of roots, then enter the entry in the set L (Flowchart 1).

6 Experimental Results

For our experimental purpose, the dataset comprises of 250 nouns, 250 verbs and 250 adjectives of Gujarati language were prepared. Input for the Morph analyzer is an inflected word and output is root word. We evaluate the result generated from the system with genuine root word in the dataset. Analysis of the result is based on the number of words mapped to the exact paradigm, number of words not mapped with the exact paradigm, number of words mapped with multiple paradigms and which category gives a good result from the system (Fig. 6).



Flowchart 1 Flowchart of the system

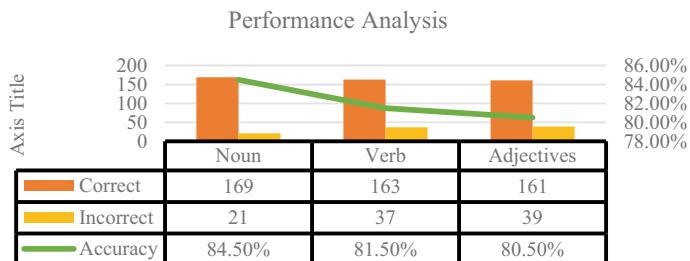


Fig. 6 Result analysis

7 Conclusion

The morphological analyzer is divided into three major modules: (1) Dictionary Creation, (2) Root Table Creation and (3) Morphological Analysis. A user-friendly interface has been created and stores all commonly used word forms for all Gujarati root words in its database. The only case in which the system might give wrong output is when all the forms of the input word are not there in the word form list.

References

1. Singh, P.D., Kore, A., Sugandhi, R., Arya, G., Jadhav, S.: Hindi morphological analysis and inflection generator for English to Hindi translation. *Int. J. Eng. Innovat. Technol. (IJEIT)* (2013)
2. Singh, H., Kumar, P.: Analysis of Noun, Pronoun and Adjective Morphology for NLization of Punjabi with EUGENE (2013)
3. Sheth, J., Patel, B.: Dhiya: a stemmer for morphological level analysis of Gujarati language. In: 2014 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), pp. 151–154. IEEE (2014, February)
4. Kapadia Utkarsh, N., Desai Apurva, A.: Morphological rule set and lexicon of Gujarati grammar: a linguistics approach. *VNSGU J. Sci. Technol.* 127–133 (2015, July)
5. Joseph, J., Anto, B.: Rule-based morphological analyzer for Malayalam nouns: computational analysis of Malayalam linguistics. *Int. J. Innovative Res. Comput. Commun. Eng.* (2015, October)
6. Sultana, A.: Description of Verb Morphology in Colloquial Bangla (2016)
7. Prathibha, R.J., Padma, M.C.: Morphological Analyzer for Kannada Inflectional Words Using Hybrid Approach (2016)
8. Kapadia, U., Desai, A.: Rule-based Gujarati morphological analyzer. *Int. J. Comput. Sci. Issues (IJCSI)* **14**(2), 30 (2017)

An Approach to Secure Collaborative Recommender System Using Artificial Intelligence, Deep Learning, and Blockchain



Monika Arora, Akanksha Bansal Chopra and Veer Sain Dixit

Abstract This paper aims at highlighting the increasing role of artificial intelligence in business and making familiar its various aspects vis-a-vis its immediate requirement in the present Indian business scenario. The paper takes into account the aspects of blockchain and deep learning components with regard to the business as future of artificial intelligence in business scenario. The study also includes the benefits and challenges of the use of artificial intelligence in business with influence of blockchain and deep learning. The relation between blockchain and deep learning and artificial intelligence has been discussed in this paper. The research collates findings from the use and implementation of components of blockchain and deep learning. The model is recommended in regard to future of artificial intelligence in business. The algorithm is written used for the implementation of artificial intelligence in business. The study concludes with the observations, future, and recommendations with respect to artificial intelligence in business with the implementation of blockchain and deep learning.

Keywords AI · Deep learning · Blockchain · Recommender system · Security

1 Introduction

The disruptive technologies are artificial intelligence and blockchain which are important these days. Every business is aiming these technologies in the support of business-to-business environment. Artificial intelligence (AI) is an area of computer science which is meant to replicate human intelligence, such as visual perception,

M. Arora (✉)

Department of IT/Operations, Apeejay School of Management, Dwarka, New Delhi, India
e-mail: marora.asm@gmail.com

A. B. Chopra

Department of Computer Science, SPM College, University of Delhi, New Delhi, India

V. S. Dixit

Department of Computer Science, ARSD College, University of Delhi, New Delhi, India

speech recognition, decision-making, or translation with the creation of intelligent machines that can undertake all such humanly acts. Artificial intelligence is thus an effort to use computers to perform those tasks that are traditionally required a human being to do. AI enables computers to learn from the stored data to use intelligently data structures and evolve procedures to the correct use of the data. Artificial intelligence can also be used in marketing where it will leverage the customer data and AI concepts like machine learning to anticipate the customer's next move and improve their journey and experience [1]. The evolution of big data and their advanced analytic solutions have changed the entire business objectives and their working. The use of machine learning has made it possible for marketers to build a clear picture of their target audiences than ever before, and in this entire process the application of artificial intelligence (AI). While implementing the AI, the blockchain helps you to give you support in building the infrastructure. Blockchain is also the growing area in the field of business. It is increasing in every area of business. It is considered to be a new Internet. Many organizations are already exploring the possibilities of the blockchain. The objective of blockchain in any organization is to develop a settlement system that processes transactions in (near) real time instead of days. There is a huge shift where organizations are shifting from e-commerce applications to blockchain-based applications. The possibilities of the use of blockchain are gigantic, and it seems that almost all the industries can use the blockchain transaction in their businesses. The blockchain can be fitted to any industry. Many business and industry will soon be disrupted by blockchain. A big fancy word "Blockchain technology" describes the recording of every event in a database. The blockchain is a database, where once data are added, it cannot be removed or altered. It contains a verifiable record history. Blockchain also offers improved transparency and veracity. All transactions carried out on the network are available to view publicly [2]. However, only certain information is visible. When the transaction happens between any two parties, the exchange of data is visible to all the parties who are connected. This will add a layer of privacy onto an otherwise transparent platform after implementation of blockchain. There is a requirement of digging into the deep for learning. This area is renowned as deep learning (DL). DL is a new era in the area of machine learning. ML is associated with ideas and neural network designed them to initiate the decision-making. Deep stands for anything extending inside that may be by means of layers [3]. Typically, learning means the acquisition of knowledge or skills through study, experience, etc. Learning can be in any form of supervised, partially supervised, or unsupervised. It is a subset of machine learning (ML) in artificial intelligence (AI). DL is also called hierarchical learning or deep structured learning.

In this paper, Sect. 1 gives introduction about artificial intelligence, blockchain, and deep learning. In Sect. 2, background and motivation behind the study were covered in Sect. 2. Section 3 highlights the related works, and Sect. 4 throws light on relationships of AI, deep learning, blockchain, and recommender systems. The section also discusses the proposed model. Section 5 discusses data analysis and interpretation. In Sect. 6, conclusion of the study has been made with future used in business.

2 Background and Motivation

In today's era of computer science, artificial intelligence and Blockchain technologies have emerged in handholding solutions. Experts believe that these technologies in combination will contribute to business implications in near future [4]. Blockchain technology first came into scene in 2008. Satoshi Nakamoto is the person who introduced blockchain technology by name "Bitcoin". This technology can track transactions between users held in public ledgers. It has been now recognized in various commercial applications. Blockchain is a system where digital information is stored in encrypted form, which in turn helps in creating of highly secured database, even in public domain. Artificial intelligence is another field of computer science which is growing very fast. It is the study where machines are expected to think and act like humans. There are many projects where working is expected as like humans and use of human intelligence is demanded. This can be achieved using artificial intelligence, machine learning, and deep learning. The integration of these technologies is beneficial as artificial intelligence is decentralized and has many independent nodes [5].

The developers of artificial intelligence also do not predict its future actions or its way of thinking [6]. The technology is capable to analyze huge amounts of data which may be beyond our imagination [7]. In blockchain technology, every piece of information is recorded securely and is kept on every computer present in the network. To manage such a huge size of data in our daily lives is another challenge to introduce blockchain [8]. With passage of time, there is an increase in the number of blocks, making the chain heavier. Optimized algorithms can be used to store data in blockchains. The dataset is split into smaller sub-datasets. These sub-datasets are then analyzed as separate units. The results of sub-datasets are finally integrated to achieve the results [9].

Related Work

Authors reviewed and presented some descriptive statistics and discussed major advancements and shortcomings of the most common recommendation concepts and approaches [10]. A recommender system is proposed that takes into account supply chain constraints regarding the availability of a product while nudging the customers to purchase it [11]. The research provides information about trends in recommender systems and future direction on recommender systems [12].

3 The Relationships

3.1 AI and Deep Learning

Deep learning is a budding new research area of machine learning. Machine learning applies THE core ideas of artificial intelligence. The major emphasizes are given on

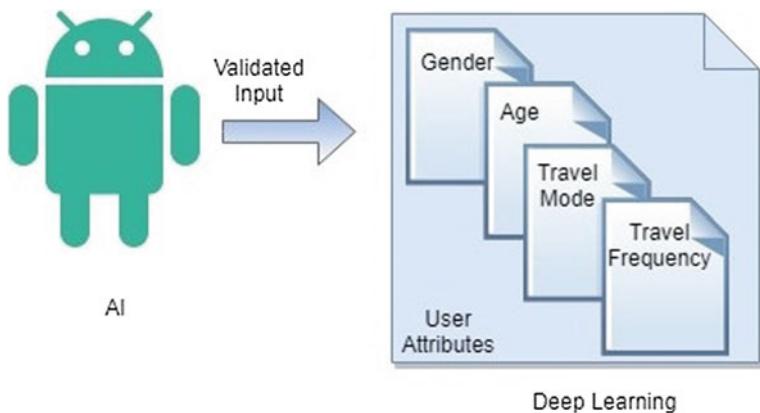


Fig. 1 Relation between AI and deep learning

solving real-world problems using neural networks. It contributes to own decision-making. Deep means extending far down. It means to study number of layers. Learning means the gaining of knowledge. It can be achieved through study, experience or experiment, etc. Deep learning is a learning of data and representing that data, based on a deeper knowledge and study. It is also called “Hierarchical Learning” or “Deep Structured Learning” [13]. Deep learning is equivalent to study the hierarchical representations. Learning can be supervised or unsupervised [5]. It is a subset of machine learning in artificial intelligence, and hence can be connected directly with AI as shown in Fig. 1.

3.2 Deep Learning and Blockchain

The relationship between the blockchain, deep learning, and artificial intelligence becomes stronger day by day. This relationship becomes very powerful and gives strength in the area of security. The future lies in these approaches. There is a tremendous increase in the use of AI applied with deep learning and blockchain, in business, in global scenario. Due to the emergent use of big data in analytics, it is a necessity to use innovative use of AI, blockchain, and deep learning as shown in Fig. 2.

3.3 Blockchain and Recommender System

A recommender system generates meaningful recommendations for items of interest to users. There are two types of approaches for recommender system: collaborative filtering [14] and content-based filtering [15]. Content-based filtering matches knowl-

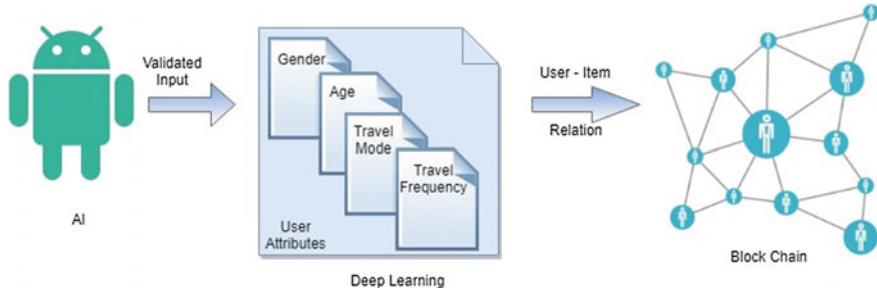


Fig. 2 Relation between AI and blockchain

edge of a single user with the knowledge of available items, whereas collaborative filtering generates recommendations based on ratings (or predictions) given by users. There is another approach called Hybrid [16] which combines both collaborative- and content-based approaches together.

In a content-based system, since a single user is involved, it does not require sharing of personal with a company. Hence, a secure recommender system is achievable [17]. In collaborative filtering system, collection of customers is involved to generate one single recommendation. As a result, companies maintain huge databases for user data. This is being done by the use of blockchain. A blockchain is an open, immutable, append-only transaction log replicated among a network of nodes [17]. In this approach, each block references its predecessor and every new block assures proof-of-work so as to get acknowledged by other participants in the network as shown in Fig. 3. This approach was first used in Bitcoin. We propose the same approach for generating predictions in recommender system, as discussed in later section.

3.4 Flow of Data Between Four Technologies

The above-discussed relationships between AI, deep Learning, blockchain, and recommender system develops another approach where data can flow between these four as described in figure below. The data input by user first enter into AI system. This can be considered a transparent layer where non-genuine user may fail to enter into the system. Only filtered data input would flow for deep learning. In this phase, hierarchical study is done, which further confirms the genuinity of user. The data in transaction verify itself and get acknowledged by other users present in the network. This is how blockchain starts forming. Once blockchains are formed, clustering algorithm may be applied to collect similar recommendations given by the users and also clustering of users with similar taste. Finally, using these predictions, recommendation generates final predictions (or ratings) for items of interest by the users as shown in Fig. 4.

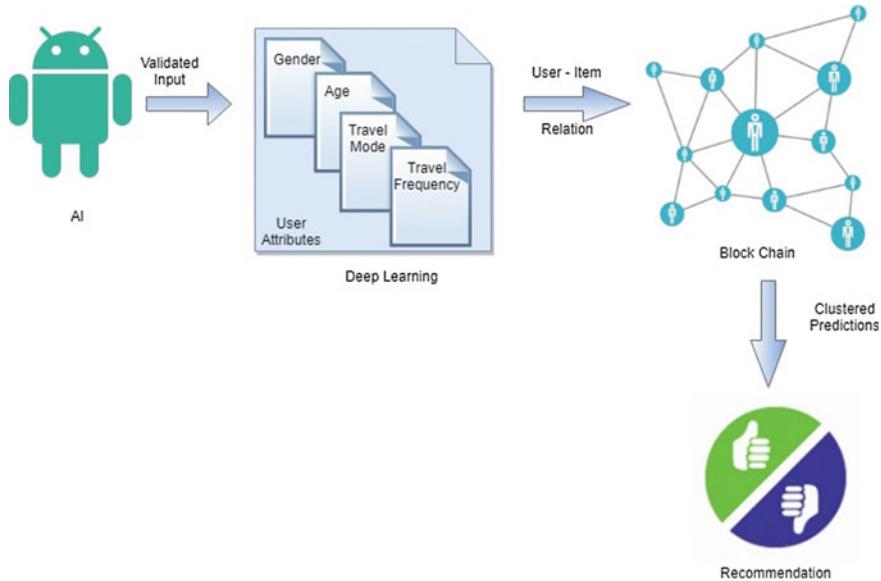
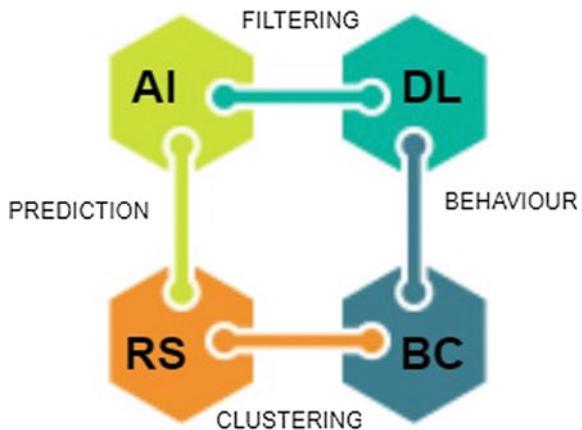


Fig. 3 Relation between AI and recommender systems

Fig. 4 Flow of data



3.5 Proposed Model

We propose a model to recommend predictions (or ratings) using three approaches—Artificial intelligence, deep learning, and blockchain. This model reduces privacy and security threat. In collaborative recommender systems, security is a high state of concern as personal data of user is required and stored by the companies [18]. The data require to be protected with primary concern, not get to be leaked or hacked. Our model serves as a solution to this problem.

We have divided our model into five stages—Stage 1: AI; Stage 2: Deep learning; Stage 3: Blockchain; Stage 4: Clustering; and Stage 5: Recommender mainframe. Stages 1–3 work under data layer, wherein stage 4 under service layer; and finally stage 5 in application layer.

In stage 1, users input their personal details and ratings (or predictions) for items of interest. The data are received and stored on the server on web. This is the transparent layer. The data flow into data layer. AI approach is used to check for any robotic interference, as discussed in our previous work. This filters the non-genuine and fake identities. Genuine data now flow to stage 2, for deep learning. In this stage, hierarchical study is executed for every individual user entry. If any fake or irrelevant or suspicious data are found, the corresponding user id and its entries are discarded. This once again filters non-genuine data. The left data are passed to stage 3 for blockchain formation. As like in Bitcoin, blocks of users are created. This helps to accommodate all data without revealing individual's identity as shown in Fig. 5.

At this stage, security is taken care off. Now, the data from blocks are picked up and clusters of “ratings of items” and “users with similar tastes” are formed. The most common and famous, k-means clustering algorithm may be used for this purpose. From the formed clusters, predictions (or ratings) for items of interest may easily be generated using recommender mainframe.

4 Data Analysis and Interpretation

We tried to implement our model using a small dataset. We conducted a survey and e-data were collected as nonparametric random sampling. There were 103 dataset genuine values. There were 54% of male and 46% of female in the sample dataset. Figure 6 depicts gender-wise distribution (Fig. 7).

The age-wise distribution of data is shown in Fig. 6. There were 61% people who belonged to age group 20–25 years, 19% in 30–35 years of age group, 10% in 35–40 years of their age, 8% in 40–45, and 2% in 25–30 years of age group. Figure 8 depicts the mode of travel in NCR region is primarily done through metro. As Fig. 8 interprets, out of 103 participants, 49 participants travel by Delhi Metro, 23 by their own car, 9 by taxi, 8 by two-wheeler, and 7 by each autorickshaw and bus.

In Table 1, frequency of travel is as follows: 23% of the participants travel once in a month, 18% of participants travel twice a day and twice a week 16% of participants travel once a week, and 13% travel once in a day. We have used above information for deep learning. Though we have used more number of parameters, limited information is depicted and highlighted in this paper. We then created blocks of this data, clustered, and generated predictions (in scale of 1–10), as shown in Table 2.

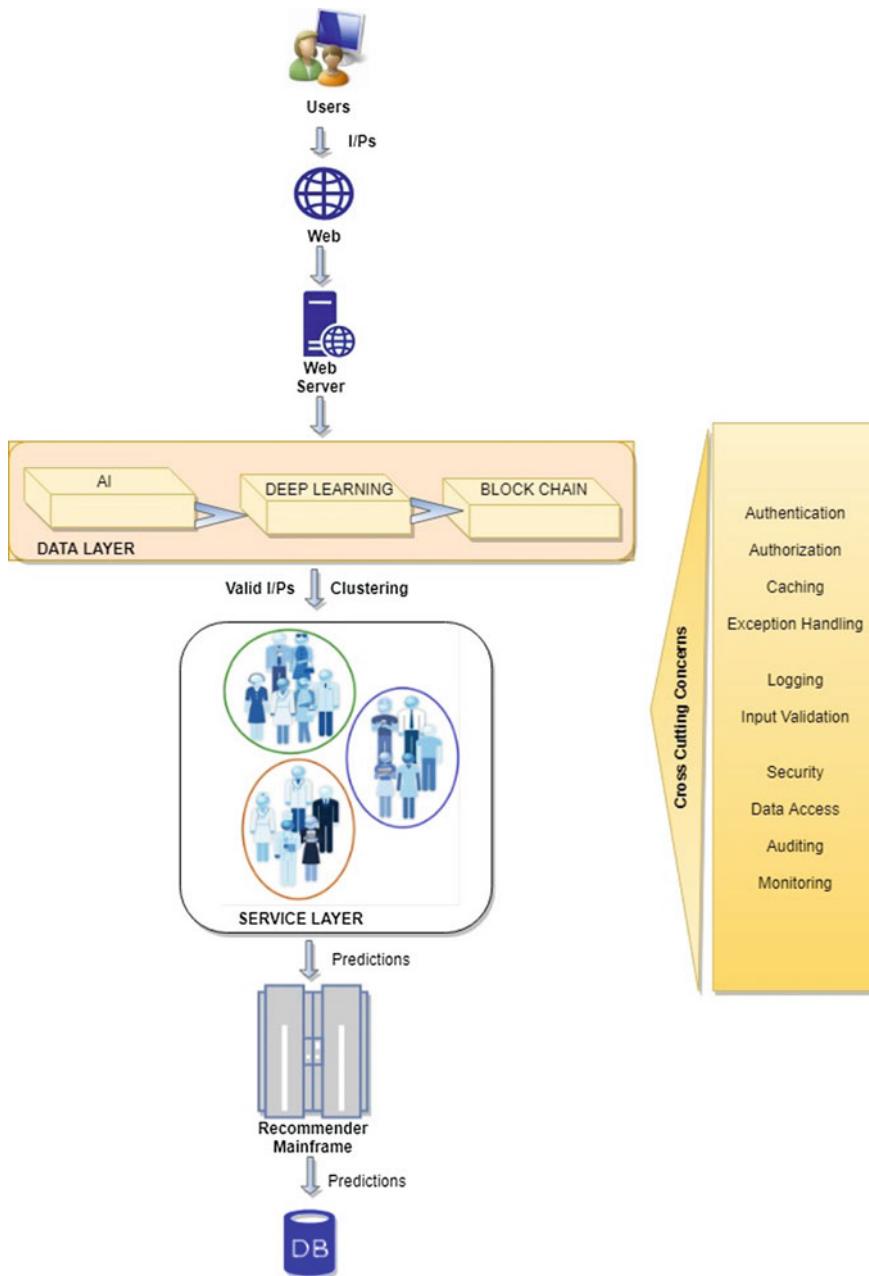


Fig. 5 Proposed model

Fig. 6 Gender-wise distribution

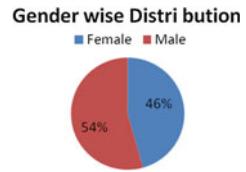


Fig. 7 Age-wise distribution

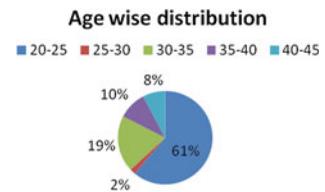


Fig. 8 Mode of travel in NCR

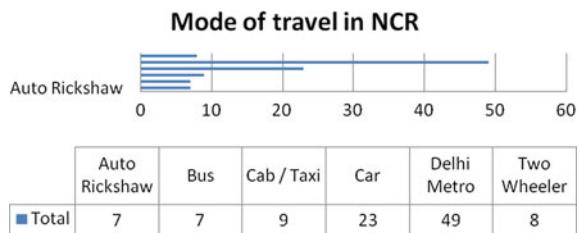


Table 1 Frequency of travel

Frequency of travel	No.	%age
Lesser	12	12
Once a day	13	13
Once a month	24	23
Once a week	16	16
Twice a day	19	18
Twice a week	19	18
	103	100

5 Validation Using Set Equations Modeling

The purpose of the study was to test whether a collaborative recommender could be made more secure by using and combining three approaches—Artificial intelligence, blockchain, and Deep Learning. A number of citations were considered to propose the model (Fig. 5). For analysis, a survey was conducted and data for various modes of travel by a number of users (with different parameters such as gender, age, travel frequency, and travel mode) was collected. Considering various citations, blockchain was prepared, and deep learning approach was used. Overall, a number of users were clustered depending upon the liking of their mode of travel, considering

Table 2 Ratings generated by recommender system

Mode of transport	Ratings	Rating stars
Autorickshaw	4.1	★★★ ★★★☆
Bus	4.0	★★★★
Car	7.1	★★★★★ ★★★★★
Cab/taxi	5.6	★★★★ ★★★★
Delhi Metro	8.5	★★★★★★ ★★★★★☆
Two-wheelers	5.4	★★★★ ★★★★

various parameters. Finally, using Jaccard similarity coefficient, recommendations are generated and depicted in Table 2.

To calculate the similarity index, variation of the Jaccard similarity coefficient is used. The formula compares similarity for the two sets with a range from 0 to 100%. The higher is the percentage, the more is the similar two populations. Mathematically,

$$\text{Jaccard Index} = (\text{the number in both sets}) / (\text{the number in either set}) * 100$$

In notation,

$$J(X, Y) = |X \cap Y| \div |X \cup Y|$$

Each user will have two sets: a set of mode of travel liked by the user (L) and a set of mode of travel disliked by the user (D). Each mode of travel will also have two sets associated with it: a set of users who liked a mode of travel (M_L) and a set of users who disliked a mode of travel (M_D).

Users' similarity index based only on the set of their liked mode of travel is defined using the Jaccard similarity coefficient formula:

$$S(U_1, U_2) = |L_1 \cap L_2| \div |L_1 \cup L_2| \quad (1)$$

Here, U_1 and U_2 are the two users, and L_1 and L_2 are the sets of mode of travel that U_1 and U_2 have liked, respectively. If two users liking the same mode of travel are similar, then two users disliking the same mode of travel should also be similar. The equation is modified as

$$S(U_1, U_2) = (|L_1 \cap L_2| + |D_1 \cap D_2|) \div |L_1 \cup L_2 \cup D_1 \cup D_2| \quad (2)$$

Adding the number of common dislikes in the numerator of Eq. (2),

$$S(U_1, U_2) = (|L_1 \cap L_2| + |D_1 \cap D_2| - |L_1 \cap D_2| - |L_2 \cap D_1|) \div |L_1 \cup L_2 \cup D_1 \cup D_2| \quad (3)$$

Subtracting the number of conflicting likes and dislikes of mode of travel of the two users from the number of their common likes and dislikes. Two users having identical tastes will have a similarity index of 100%, while two users having entirely conflicting tastes will have a similarity index of 0%.

If $P(U, M)$ is the probability of a user U liking the mode of travel M , S_L and S_D are the sum of the similarity indices of user U with all the users who have liked or disliked the mode of travel M , respectively. $|M_L| + |M_D|$ represents the total number of users who have liked or disliked the mode of travel M .

$$P(U, M) = (S_L - S_D) \div (|M_L| + |M_D|) \quad (4)$$

Equations (1)–(4) are used to generate the recommendations for mode of travel, depicted in Table 2.

6 Conclusion and Future

In this paper, we present AI and deep learning, a robust platform based on blockchain technology for a secured collaborative recommender system. Initially, we started with establishing relationships and relevance between AI, deep learning, blockchain, and recommender system. We took a training set, collected through a survey, to check the feasibility of our study. We desired to use these technologies\approaches to propose a model to develop a secured collaborative recommender system. While conducting our experimental study, we used AI approach as the first level of security. The relevant and genuine data are passed to next level, i.e., deep learning or for hierarchical study. To preserve ethics, we maintained privacy of participants of training set, by not highlighting all details of the dataset. With both AI and blockchain technologies together in one place, you can easily control how your data can be used for the dataset you own. For various industries where privacy of data is important, this technology will come in handy. With both the technologies currently trending in the industry, the combination of these two technologies promises to provide greater benefits in the coming future.

At this level, again data found with fake entries are eliminated. Next, blocks of obtained data from hierarchical study are formed. This step helps to hide individuals' details, which is a major portion of security concern for any company. Finally, clustering of items and users with similar taste is done, using k-means clustering algorithm, and ratings for items are obtained. With AI, deep learning, and blockchain technology together in one place, you can easily control how your data can be used for the dataset you own. For various industries where privacy of data is important, these technologies combined with recommender system will come in handy. With both the

technologies currently trending in the industry, the combination of these technologies promises to provide greater benefits in the coming future. We also intend to redefine the security problem and find more solutions in future work.

References

1. Arora, M., Datta, A.: Artificial intelligence and consumer behaviour in marketing. *Trends Mark. Contemp. Issues*, pp. 42–47 (2018)
2. Arora, M., Arora, A.: Digital information tracking framework using block chain. *J. Supply Chain Manag. Syst. (UGC Approved)* **7**(2), 1–7, ISSN: 2277-1387 (2018)
3. Lotter, W., Kreiman, G., Cox, D.: Deep predictive coding networks for video prediction and unsupervised learning. arXiv preprint [arXiv:1605.08104](https://arxiv.org/abs/1605.08104) (2016)
4. Baker, J., Steiner, J.: Blockchain: The Solution for Transparency in Product Supply Chains. Provenance, London, UK (2015)
5. Rosoff, M.: IBM and Salesforce shake hands on artificial intelligence. CNBC. 6 Mar 2017
6. Selyukh, A.: Tech giants team up to tackle the ethics of artificial intelligence. Npr.org. 28 Sept 2016
7. Shmueli, G., Koppius, O.R.: Predictive analytics in information systems research. *MIS Q.* **35**(3), 553–572 (2011)
8. The 2016 AI Recap: Startups See Record High In Deals And Funding. Cbinsights.com. 19 Jan 2017
9. Paré, G., Trudel, M.-C., Jaana, M., Kitsiou, S.: Synthesizing information systems knowledge: a typology of literature reviews. *Inf. Manag.* **52**, 183–199 (2015)
10. Beel, J., Langer, S., Breitinger, C.: Research-paper recommender systems: a literature survey. *Int. J. Digit. Libr.* **17**(4), 305 (2016)
11. Zhang, N., Kannan, K., N., Shanthikumar, J.G.: A recommender system to nudge customers in a capacity constrained supply chain. Available at SSRN: <https://ssrn.com/abstract=2588523> or <http://dx.doi.org/10.2139/ssrn.258852> (2015)
12. Park, D.H., Kim, H.K., Choi, I.Y., Kim, J.K.: A literature review and classification of recommender systems research. *Expert Systems with Applications* (39: 1) pp. 10059–10072 (2012)
13. Yann, L., Yoshua, B., Geoffrey, H.: Deep learning. *Nat. Int. J. Sci.* **521**, 436–444 (28 May 2015) (2015)
14. Goldberg, D., Nichols, D., Oki, B.M., Terry, D.: Using collaborative filtering to weave an information tapestry. *Commun. ACM* **35**(12), 61–70 (1992)
15. Pazzani, M.J.: A framework for collaborative, content-based and demographic filtering. *Artif. Intell. Rev.* **13**(5), 393–408 (1999)
16. Burke, R.: Hybrid recommender systems: survey and experiments. *User Model. User-Adapt. Interact.* **12**, 331–370 (2002)
17. Frey, R., Worner, D., Ilic, A.: Collaborative filtering on the blockchain: a secure recommender system for e-commerce. In: Twenty-second Americas Conference on Information Systems, San Diego. Available from: https://www.researchgate.net/publication/306119856_Collaborative_Filtering_on_the_Blockchain_A_Secure_Recommender_System_for_e-Commerce (2016)
18. Adomavicius, G., Bockstedt, J., Curley, S., Zhang, J.: De-biasing user preference ratings in recommender systems. In: Joint Workshop on Interfaces and Human Decision Making in Recommender Systems (2014)
19. Duhigg, C.: How Companies Learn Your Secrets. Available from <http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=all> (2012)
20. Jindal, H., Singh, S.K.: A Hybrid recommendation system for cold-start problem using online commercial dataset. *Int. J. Comput. Eng. Appl.* **VII**(I) (2014)
21. Hosanagar, K., Fleder, D., Lee, D., Buja, A.: Will the global village fracture into tribes: recommender systems and their effects on consumers. Working Paper (2013)

22. Gujral, M., Chandra, S.: Knowledge synthesis of recommendation systems—finding expert recommendations for cuisines. *Int. J. Technol. Enhanc. Eng. Res.* **3**(5), 88–94 (2015)
23. Ruth, Y., Ilaria, B., Brad, S.: The future of artificial intelligence in consumer experience, 3–10 (2017)

Performance Analysis of NSL-KDD Dataset Using Classification Algorithms with Different Feature Selection Algorithms and Supervised Filter Discretization



Shailesh Singh Panwar and Y. P. Raiwani

Abstract Naive Bayes and Bayes Net are critical classification methods in data mining classification and have build up being important software tools for the classification, the description and generalization of information. All classification algorithms are open sources, which are implemented on Java (C4.5 algorithms) for WEKA software tool. In this paper, we exhibit the strategy for increasing the performance of Naive Bayes and Bayes Net algorithms with supervised filter discretization after we applied feature selection techniques. We have used the supervised filter discretization on two classification algorithms and compared the result with and without discretization. The outcomes acquired from the experiment showed significant improvement over the existing classification algorithms.

Keywords Naive Bayes · Bayes net · WEKA · NSL-KDD dataset · Preprocessing · Discretization · Feature selection algorithms

S. S. Panwar (✉) · Y. P. Raiwani
H.N.B. Garhwal University, Srinagar, Garhwal, Uttarakhand, India
e-mail: Shaileshpanwar23@gmail.com

Y. P. Raiwani
e-mail: yp_raiwani@yahoo.com

1 Introduction

Data mining is the process of fetching valuable information from huge amount of data stored in database, data warehouse, or other repositories. It is the computational process of finding patterns from large data sets involving methods at the intersection of artificial intelligence, machine learning, and database systems. In data mining applications, we can use several types of parameters to examine the data [1–3].

In data mining, real-world problems often involve a large number of features. In these features, many of them are redundant or irrelevant, so all features are not important. These redundant features reduce the performance of algorithms. Feature selection process is used to solve this problem by choosing a small subset of similar features from the original huge set of features. By eliminating irrelevant and redundant features, the performance of learn models is enhanced such as speed of learning process is increased, and dimension of data is reduced [4].

Classification is a data mining technique, which is used for implementing various projects. It predicts the value of a categorical attribute based on the values of other predicting attributes. A search algorithm is used to induce a classifier from a set of correctly classified data instances known as the training set. The quality of the classifier is measured by another set of correctly classified data instances, called the testing set [5].

Naive Bayes and Bayes Net are the simplest probabilistic classifiers. These classifiers are used in various real-world applications and assume that all features are independent of each other in the given class. In the learning process of these classifiers with the particular model, class probabilities and conditional probabilities are evaluated using training data, and then new observations are classified by using values of these probabilities [6].

This paper is organized as follows:

Phase 2 presents our proposed approach. Phase 3 gives a brief description of methodologies, discretization, and feature reduction techniques. Phase 4 analyzes the results and improves the performance of classification algorithms and after that finds the best techniques. At last, Phase 5 gives conclusions.

2 Proposed Approach

Step 1. Define problem and select the dataset.

Step 2. Apply data preprocessing on dataset.

Step 3. Import dataset into software tool.

Step 4. Apply feature reduction techniques.

Step 5. Used supervised filter discretization on dataset.

Step 6. Select single classification algorithms

Step 7. Apply classification algorithms on dataset.

Step 8. Based on measure, analyze the performance of classification algorithms.

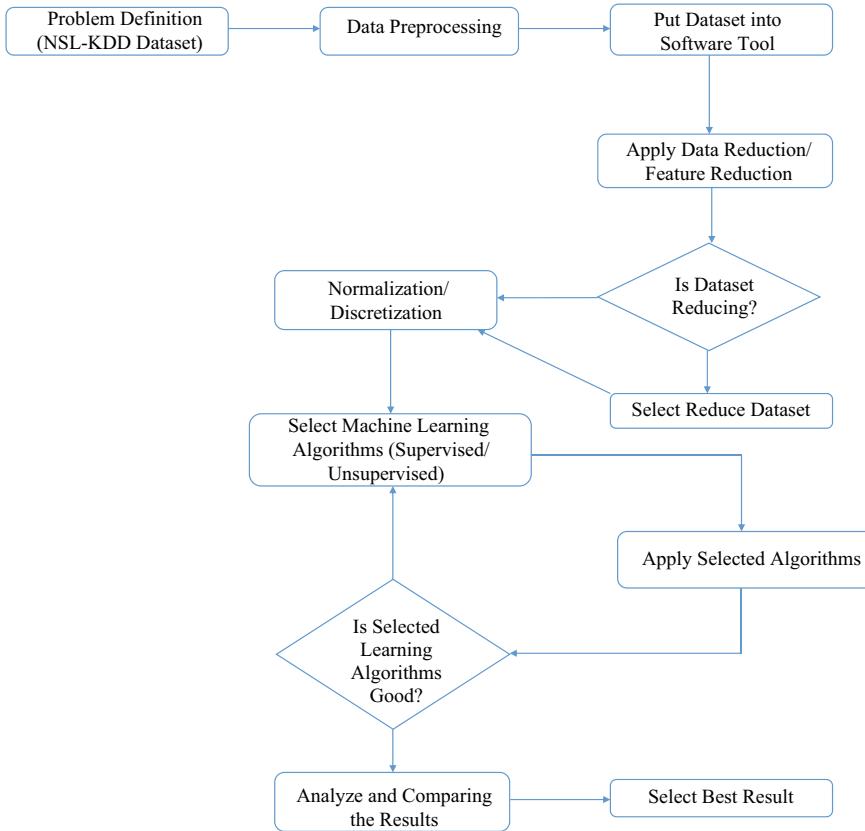


Fig. 1 Proposed approach

Step 9. If the results are as per measure, consider the classification algorithms for further processing.

Step 10. Go to step 6 and repeat from steps 6 to 9 other classification algorithms.

Step 11. Analyzed the results of classification algorithms based on performance measures and evaluate the best result (Fig. 1).

3 Feature Reduction and Classification Algorithms

3.1 Correlation Feature Selection (CFS)

Feature selection is a process of choosing a subset of the relevant attribute selected in a large number of basic attribute of a particular dataset by applying unique assessment

standards to enhance the pleasant of classification, while the dimension of the data reduces. It is used to evaluate the subset of features based on the well-suited subsets, which have highly correlated facilities with classification, are still unrelated to each other [7, 8].

3.2 Classifier Subset Evaluator

- It evaluates the specialty subset on training data or a separate hold-out test set.
- It uses a classification to estimate the eligibility of a set of attributes.
- With whom the classification algorithms perform well, it considers subsets of those tasks.

3.3 Naive Bayes Classifier

Naive Bayes Classifiers is a family of “probabilistic classification”, which implements with powerful independence hypothesis between the features. It is particularly scalable; require a number of linear parameters inside the wide variety of irregular capabilities in a learning problem. Naive Bayes is an easy technique for building classifiers: sample and pattern that gives grandiosity labels to difficult situation time, which are shown as feature value vectors, where the labels are exhausted from some limited set. This is not the only algorithm for such classification, but the group of algorithm is entirely based on a general preaching. All Naive Bayes classifiers anticipate that the significance of a specific attribute is independent of the significance of any other attribute, given the class variable [9].

3.4 Bayes Net

In machine learning, Bayes Net classification techniques are a group of belief networks. Bayesian networks are the probabilistic networks. This algorithm uses Learn Bayesian nets functions. Bayes Net algorithm uses various search algorithms and quality measures. It provides data structure and facilities common to Bayes network learning algorithms [10].

3.5 *Preprocessing and Discretization*

Records and information are generally common in assorted layout: nominal, continuous and discrete. Discrete and continuous statistics having orders among values are ordinal records types. But there is no order in the nominal values. Preprocessing is done by supervised filter discretization technique. Discretization is categorized into two different categories, supervised and unsupervised discretization. Unsupervised Discretization is always preferred when no class information is available. The EWB algorithm is the simplest one and it determines the minimum and maximum values of the discretized attribute in order to determine its range [11]. Regulated Discretization procedures as the name recommends considers the class data before making subgroups. Administered strategies are predominantly in light of Fayyad and Irani [12] or Kononenko [13] calculations.

4 Experiments and Results

4.1 WEKA Tool

In this paper, we have used the WEKA Software tool to investigate and analyze the NSL-KDD dataset with machine learning algorithms. WEKA is an open-source GUI application, which is referred to as the Waikato Environment for Knowledge Learning. The University of Waikato in New Zealand developed the WEKA software tool, which identifies the data from lager amount records that is collected from different domains. It helps on several data mining and machine learning applications along with preprocessing, clustering classification, regression, feature selection, and visualization regression [14].

4.2 NSL-KDD Dataset

NSL-KDD dataset was used to solve some of the implied issues of KDD 99 dataset. There are no duplicate data in the test set proposed in the NSL-KDD dataset. Therefore, the performance of newcomers is not biased through the methods, which have a better identification rate on common data. This dataset contains variety of attributes, which can be supportive for measure the attacks. NSL-KDD dataset have 22,544 instances at dataset (KDD Test) and 125,973 instances for training dataset (KDD Train) [15, 16] (Fig. 2).

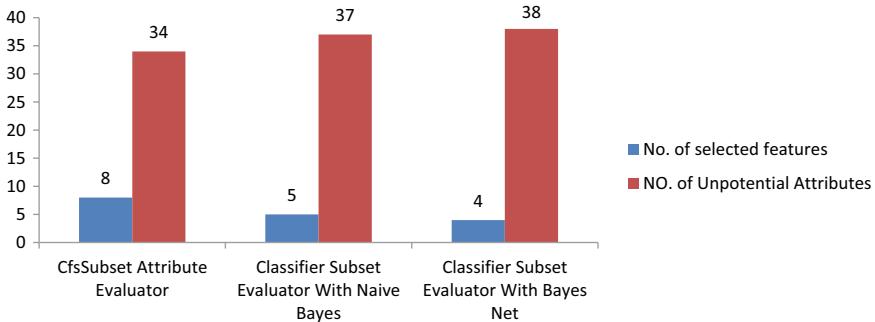


Fig. 2 No. of selected attributes after applying feature reduction algorithm

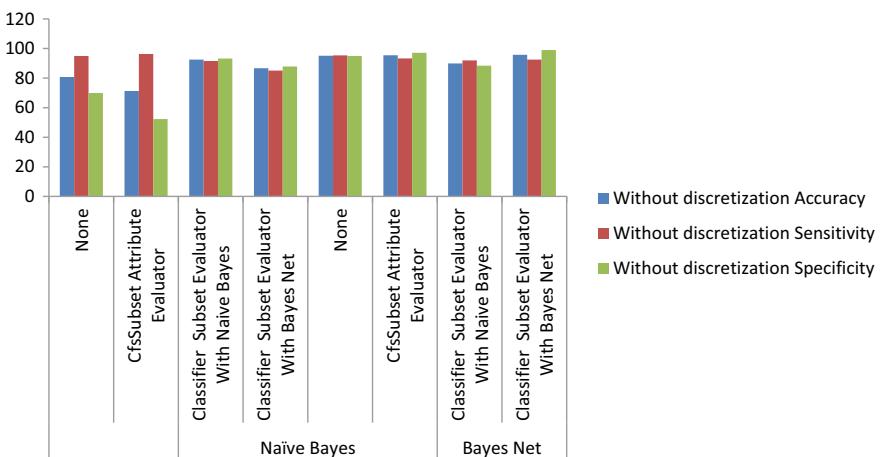


Fig. 3 Analyzing best algorithm for testing data without discretization

4.3 Performance Evaluation

WEKA software tool is applied in both training and testing datasets (NSL-KDD datasets), and finds out the accuracies by Naïve Bayes, Bayes algorithms without supervised discretization and with supervised discretization. The accuracy, sensitivity, and specificity are obtained as shown in Figs. 3, 4, 5, and 6. The final result indicates that supervised discretization with feature selection algorithms has improved the overall performance of machine learning algorithms for both training and as well as testing data.

Table 1 shows that after applying feature selection algorithms on the complete dataset, the selected data and selected properties to reduce the dataset are included. Each algorithm does not have a different number of properties based on their evaluation criteria. Now, the main task is to find out which classification algorithms give

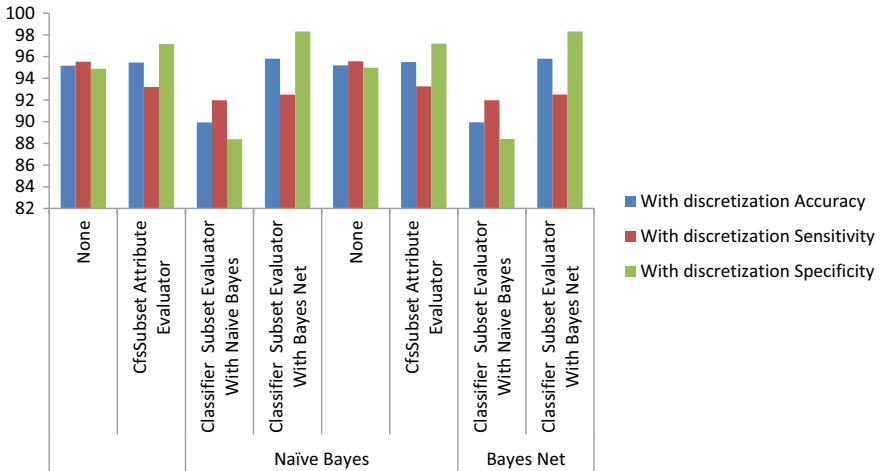


Fig. 4 Analyzing best algorithm for testing data with discretization

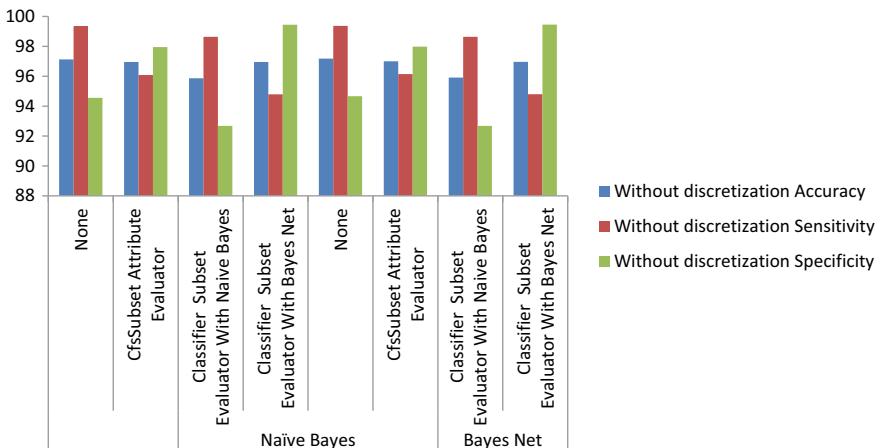


Fig. 5 Analyzing best algorithm for training data without discretization

better results for feature selection on NSL-KDD datasets. For this purpose, we have implemented two famous classification algorithms, Fools Bayes and Bias Nets, and the data reduction algorithm of these two algorithms gives better results in terms of accuracy, sensitivity, and specificity. Figure 2 shows the feature selection algorithm output. The number of selected attributes is given in the form of input in the classification algorithm. According to Tables 3, 4, 5, 6, 7, and 8, the comparisons show that using two selected classifiers with all features and features chosen by CSF, Classifier Subset Evaluator with Naive Bayes and Bayes Net yields the following results (Table 2):

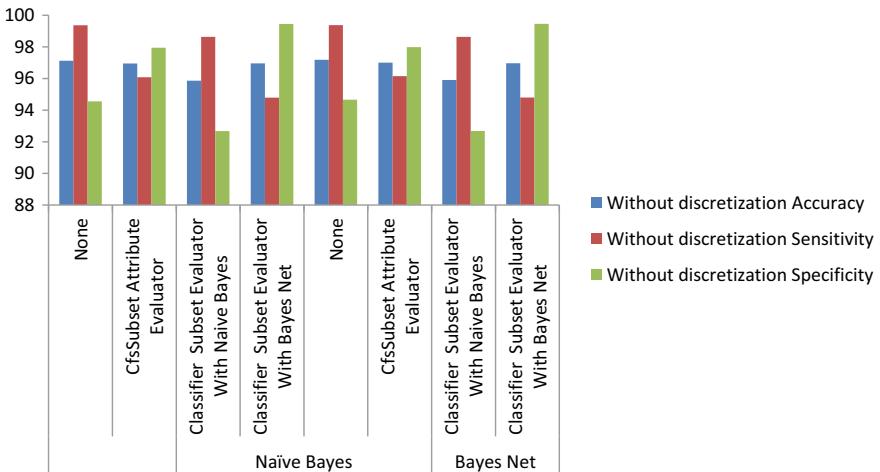


Fig. 6 Analyzing best algorithm for training data with discretization

Table 1 The selected features using CfsSubset attribute evaluator, classifier subset evaluator with Naïve Bayes, and classifier subset evaluator with Bayes net

Features selection algorithm	Selected features	No. of selected features
CfsSubset attribute evaluator	5, 6, 12, 28, 30, 31, 37, 41	8
Classifier subset evaluator with Naïve Bayes	3, 4, 23, 24, 36	5
Classifier subset evaluator with Bayes Net	3, 5, 6, 13	4

- The Naïve Bayes classifier with features selected by Classifier Subset Evaluator with Bayes Net technique gives the best performance for the test data (With discretization).
- The Bayes Net classifier with features selected by CfsSubset Attribute Evaluator technique gives the best performance for the training data (With discretization).
- The Bayes Net classifier with features selected by CfsSubset Attribute Evaluator technique gives the best performance for the training data (Without discretization).
- The Bayes Net classifier with features selected by Classifier Subset Evaluator with Bayes Net technique gives the best performance for the testing data (Without discretization).

Table 2 List of attributes

Duration	wrong_fragment	su_attempted	is_gust_login	same_srv_rate	dst_host_same_src_port_rate
Service	Urgent	num_root	srv_count	diff_srv_rate	dst_host_srv_diff_host_rate
Protocol_type	Hot	num_file_creation	count	srv_diff_host_rate	dst_host_error_rate
Flag	num_failed_login	num_shells	error_rate	dst_host_count	dst_host_srv_serro_rate
src_byte	logged_in	num_access_file	srv_error_rate	dst_host_srv_count	dst_host_error_rate
dst_byte	num_outbound_cmds	num_outbound_cmds	error_rate	dst_host_same_srv_rate	dst_host_srv_rerror_rate
Land	root_shell	is_host_login	srv_rerror_rate	dst_host_diff_srv_rate	Class

Table 3 Performance evaluation for testing data without discretization

Selection algorithms		None	CfsSubset attribute evaluator	Classifier subset evaluator with Naive Bayes	Classifier subset evaluator with Bayes net
Classifier	Metric				
Naïve Bayes	Correctly classified	18,200	21,519	20,854	19,538
	Incorrectly classified	4344	1025	1690	3006
	Kappa	0.623	0.9069	0.8472	0.7285
	Time taken	0.45	0.01	0.16	0.08
Bayes net	Correctly classified	21,446	21,529	20,277	21,593
	Incorrectly classified	1098	1015	2267	951
	Kappa	0.9009	0.9079	0.7967	0.9134
	Time taken	1.31	0.03	0.47	0.58

Table 4 Performance evaluation for testing data with discretization

Selection algorithms		None	CfsSubset attribute evaluator	Classifier subset evaluator with Naive Bayes	Classifier subset evaluator with Bayes net
Classifier	Metric				
Naïve Bayes	Correctly classified	21,452	16,067	20,274	21,599
	Incorrectly classified	1092	7477	2270	945
	Kappa	0.9015	0.4536	0.7965	0.914
	Time taken	0.09	0.17	0.02	0.01
Bayes net	Correctly classified	21,461	21,523	20,277	21,599
	Incorrectly classified	1083	1021	2267	945
	Kappa	0.9023	0.9073	0.7968	0.914
	Time taken	0.33	0.2	0.02	0.01

Table 5 Performance evaluation for training data without discretization

Selection algorithms		None	CfsSubset attribute evaluator	Classifier subset evaluator with Naive Bayes	Classifier subset evaluator with Bayes net
Classifier	Metric				
Naïve Bayes	Correctly classified	113,858	78,620	119,676	68,089
	Incorrectly classified	12,115	57,353	6297	57,884
	Kappa	0.806	0.2104	0.8994	0.0143
	Time taken	1.28	0.53	0.23	0.25
Bayes net	Correctly classified	122,409	122,122	120,796	122,113
	Incorrectly classified	3564	3851	5177	3860
	Kappa	0.943	0.9387	0.9171	0.9386
	Time taken	8.86	2.05	1.48	0.0331

Table 6 Number performance evaluation for training data with discretization

Selection algorithms		None	CfsSubset attribute evaluator	Classifier subset evaluator with Naive Bayes	Classifier subset evaluator with Bayes net
Classifier	Metric				
Naïve Bayes	Correctly classified	122,353	122,134	120,760	122,139
	Incorrectly classified	3620	3839	5213	3834
	Kappa	0.9421	0.9389	0.9165	0.9391
	Time taken	0.11	0.05	0.03	0.06
Bayes net	Correctly classified	122,422	122,197	120,762	122,151
	Incorrectly classified	3551	3776	5211	3822
	Kappa	0.9432	0.9399	0.9166	0.9392
	Time taken	0.45	0.11	0.08	0.09

Table 7 Accuracy, sensitivity, and specificity for testing data with discretization and without discretization

Classifier	Features selection algorithm	Without discretization			With discretization		
		Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
Naïve Bayes	None	80.731	94.9953	69.9368	95.1561	95.5205	94.8803
	CfsSubset attribute evaluator	71.2695	96.3237	52.31	95.4533	93.1932	97.1635
	Classifier subset evaluator with Naive Bayes	92.5035	91.5559	93.2206	89.9308	91.9781	88.3815
	Classifier subset evaluator with Bayes net	86.6661	85.1302	87.8282	95.8082	92.5130	98.3168
Bayes net	None	95.1295	95.3866	94.9349	95.1961	95.572	94.9815
	CfsSubset attribute evaluator	95.4711	93.3374	97.0856	95.4977	93.2550	97.1947
	Classifier subset evaluator with Naive Bayes	89.9441	91.9369	88.4360	89.9441	91.9781	88.4048
	Classifier subset evaluator with Bayes net	95.7816	92.4930	98.97	95.8082	92.5033	98.3090

Table 8 Accuracy, sensitivity, and specificity for training data with discretization and without discretization

Classifier	Features selection algorithm	Without discretization			With discretization		
		Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
Naïve Bayes	None	90.3829	93.64	86.6416	97.1264	99.3644	94.5556
	CfsSubset attribute evaluator	62.4102	93.98	26.1487	96.9525	96.0842	97.9498
	Classifier subset evaluator with Naive Bayes	95.0013	96.5757	93.1929	95.8618	98.6338	92.6778
	Classifier subset evaluator with Bayes net	54.0505	99.5173	98.1732	96.9565	94.7878	99.4473
Bayes net	None	97.1708	99.3540	94.6631	97.1811	99.3718	94.6648
	CfsSubset attribute evaluator	96.943	96.0708	97.9447	97.0025	96.1465	97.9856
	Classifier subset evaluator with Naive Bayes	95.8904	98.6279	92.7460	95.9086	98.6338	92.6812
	Classifier subset evaluator with Bayes net	96.9359	94.7566	99.4388	96.966	94.7987	99.4559

5 Conclusion

In the initial phase of data mining, the preprocessing process shows its advantages amid the order exactness execution experiment. In this paper, filtered supervised discretization strategy after feature reduction techniques has been used to increase the characteristic accuracy of datasets, with the continuous valuable features. In the first phase, we applied feature reduction techniques and discretized the continuous valuable features from the given datasets on dataset. In the second step, we are going to execute Naive Bayes and Bayes Net with and without supervised discretization and finally, the results have been compared. According to the results, the use of feature selection and discretization reduce data dimension, time, false alarm, and gives higher performance results. The result shows that the feature reduction and filtering supervised discretization has a bigger effect on the execution of the classification algorithms.

In future, we can perform this work on large-size dataset and can compare large No. of feature selection techniques. We can also combine the feature selection with clustering to enhance the efficiency of the algorithms.

References

1. Kantardzic, M.: Data Mining: Concepts, Models, Methods, and Algorithms. Wiley, ISBN: 0471228524 (2003)
2. Mitra, S., Acharya, T.: Data Mining Multimedia, Soft Computing, and Bioinformatics. Wiley (2003)
3. Shun, J., Malki, H.A.: Network intrusion detection system using neural networks. Fourth international conference on natural computation, Jinan, pp. 242–246 (2008)
4. Xue, B., Zhang, M., Browne W.N., Yao, X.: A survey on evolutionary computation approaches to feature selection. In: IEEE transactions on evolutionary computation, vol. 20, no. 4, pp. 606–626, Aug. 2016 (2016)
5. Espejo, P.G., Ventura, S., Herrera, F.: A survey on the application of genetic programming to classification. In: IEEE transactions on systems, man, and cybernetics, part C (Applications and reviews), vol. 40, no. 2, pp. 121–144, March 2010 (2010)
6. Taheri, S., Mammadov, M.: Learning the naive bayes classifier with optimization models. Int. J. Appl. Math. Comput. Sci. **23**(4), 787–795 (2013)
7. Panwar, S.S., Raiwani, Y.P.: Data reduction techniques to analyze NSL-KDD dataset. Int. J. Comput. Eng. Technol. **5**(10), 21–31 (2014)
8. Raiwani, Y.P., Panwar, S.S.: Research challenges and performance of clustering techniques to analyze NSL-KDD dataset. Int. J. Emerg. Trends Technol. Comput. Sci. (IJETTCS) **3**(6), 172–177 (2014)
9. <http://cis.poly.edu/~mleung/FRE7851/f07/naiveBayesianClassifier.pdf>
10. <http://web.ydu.edu.tw/~alan9956/>
11. Gama, J., Pinto, C.: Discretization from data streams: applications to histograms and data mining. In: Proceedings of the 2006 ACM Symposium on Applied Computing, SAC, New York, NY, USA, pp. 662–667 (2006)
12. Fayyad, U.M., Irani, K.B.: Multi-interval discretization of continuous valued attributes for classification learning. In: Thirteenth International Joint Conference on Artificial Intelligence, vol. 2, pp. 1022–1027. Morgan Kaufmann Publishers (1993)

13. Kononenko, I.: On biases in estimating multivalve attributes. In: 14th International Joint Conference on Artificial Intelligence, pp. 1034–1040 (1995)
14. WEKA User Manual: [Available Online] www.gtbit.org/downloads/dwdmsem6/dwdmsem6lman.pdf
15. NSL-KDD dataset: [Available Online] <http://iscx.ca/NSL-KDD/>
16. Raiwani, Y.P., Panwar, S.S.: Data reduction and neural networking algorithms to improve intrusion detection system with NSL-KDD dataset. Int. J. Emerg. Trends Technol. Comput. Sci. (IJETTCS) **4**(1), 219–225 (2015)

An Efficient Method for Testing Source Code by Test Case Reduction, Prioritization, and Prioritized Parallelization



Pradeep Udupa, S. Nithyanandam and A. Rijuvana Begum

Abstract Software Testing is the process of verifying and validating the system with the goal of detecting and eliminating the errors, and it involves validating an attribute to see that whether it generates expected and required outputs. Here, apfd, test case reduction, prioritization, and test case rank are used to prolong performance and an algorithm is developed to optimize the overall testing efficiency and to reduce the execution time by reducing number of test cases, and then performing prioritization and fault detection, and further prioritized parallelization is used to maximize performance.

1 Introduction

1.1 Testing Overview

Software Testing is the technique of detecting, exploring, and correcting the errors. It is used to ensure the software quality and completeness. Here, the goal is to minimize total test runs [1]. Because as the number of test case increases, it takes more time to test, therefore here, we try to minimize test cases then prioritize and optimize them.

Black Box Test: It takes the input and checks whether an executing program gives proper output, but in this case, how it checks whether function works properly or not, implementation details are appropriate or not, whether software behaves properly or not is not visible to the testing team as in [2].

P. Udupa (✉)

Research Scholar, Department of C.S.E, Prist University, Thanjavur, India

e-mail: pradeepumlr@gmail.com

S. Nithyanandam

Department of Computer Science and Engineering, Prist University, Vallam, Thanjavur, India

A. Rijuvana Begum

Research Guide, Department of E.C.E, Prist University, Thanjavur, India

White Box Test: In this case, the process of checking whether function works properly or not, implementation details are appropriate or not, whether software behaves properly or not is visible to the testing team. It will check whether the application works or not. Test cases are used to define required and expected output and it is used for testing against user requirements and against specific criterias to be satisfied and is used when software needs meet their certain client's objective.

1.2 Check Lists or Test Case

A test case is designed to test whether system works properly or not basically, it is a one step, or it is a number of steps, it is used to test the correct behavior or operations and features of an application. An expected result or an expected outcome. It is used for checking whether we can able to obtain our desired output as per our requirements, which is stated at the beginning as expected result.

1.3 Validation Suite

A validation suite is a group of test cases used to validate the system with respect to given constraints or requirements it used to validate whether the system gives expected output. Test cases and information on the system configuration is used during testing to ensure validation criteria and specified requirements. It is a collection of test cases with underlying test criteria's and specific outcomes to be obtained.

1.4 Testing Control Flow

It is used to test every possible path, it can be used when the number of paths are more and testing more number of paths are complex and time consuming, and it will be helpful in checking different criterias to be satisfied by the given program it includes identifying different possible paths, and testing different possible paths.

1.5 Independent Paths

It is a unique or separate path in the program, which is used to test specific criteria and different criterias to be satisfied by the program and generate test cases for each and every derived unique path by making a specific path to pass the required condition.

1.6 Why to Lessen Overall Test Suits?

1. Bigger the test cases more will be the complexity as in [2].
2. Larger the test cases more will be the probable number of errors.
3. Error tracing is to be performed for every test case.
4. Huge number of testers are needed.
5. It will take a long time.
6. It will increase the overall cost as in [3].

2 Literature Review

In the last few years, there were many publications, which discussed the concept of test case reduction. In this section, different test case reduction methods and related works are discussed. Constraint-based test data generation by Richard A. DeMillo and A. Jefferson Offutt had presented an approach to test data generation that uses control flow analysis, symbolic evaluation, and reduces the number of test cases based on criteria. Dynamic Domain Reduction (DDR) by A. Jefferson Offutt Zhenyi Jin used get split algorithm, which divides domain to reduce the overall domain range and reduces test runs and total time required to execute. It has achieved more depletion in test cases but it is less efficient and time consuming and more efforts are needed. Ping-Pong Technique: This technique selects less minimum number of test cases by arranging differently, works using heuristic method which wont ensure best result but gives good result in given time, by contrasting the set of values of goal state and set of states of achieved values and assure domain coverage. But its time consuming and more expensive technique. *Test Case Reduction Using Parallelization on Many* techniques are represented previously in literature but in our technique we used test case reduction approach, prioritization, fault detection and parallelization where *min, max, constant variable in all paths are found and later more than one test case are made to run in parallel fashion which has greatest percentage of reduction in terms of number test cases and in execution time required for achieving parallelization, debugging (Table 1)*.

3 Existing Methodology DDR Technique

A step is as follows assume that given domain is $i1(0..30)$, $j1(0..50)$, $k1(0..40)$

1. Detecting all criterias from beginning to end.
2. Evaluate split point value for range of domain and for all variable satisfying criterias.
3. Then as per split vale, we segregate it into two intervals. $Ma1 = 0$ to 15 and $16-30$ $ma2$ into $10-30$ and $31-50$ and derive the final interval by using splitting

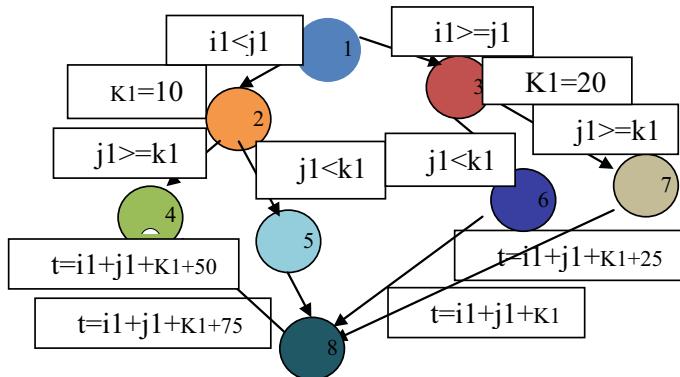
Table 1 Contrast among different depletion approaches for test cases

S no.	Test reduction technique Name	Advantage	Disadvantage
1	Constraint-based testing	It uses control flow analysis, symbolic evaluation and reduces number of test cases based on criteria	<ol style="list-style-type: none"> 1. More number of test cases 2. Time consuming 3. More expensive 4. Less efficient 5. More effort 6. No parallelization
2	DDR	Achieved more depletion in test cases	<ul style="list-style-type: none"> • More test cases compare to constraint bases • Comparatively more time consuming. • More expensive • More effort • No parallelization
3	Ping-Pong technique	Assure domain coverage and cost effective	Time consuming, expensive technique more memory, time, and efficiency is required in executing the test cases
4	Prioritize techniques no order, reverse order	Attempt to detect possible no. of errors and contribute in performance	Less efficient compared to proposed technique in terms of performance

value is ma1 0–10 and 11–30 and ma2 31–50 and ma3 is 10. So, the total test cases = $31 * 1 + 31 * 20 = 651$.

4 Proposed Methodology

Here, we first reduce the test cases by our given algorithm, then we prioritize test cases by giving rankings for test cases, and then we find APFD value which will prove that ours technique over performed. Then existing techniques, then parallelize our test cases to reduce time and cost involved. Here, first, we find number of test paths then from each path we find min, max, and constant values and will derive our reduced test cases by using steps given below, and then further execute them in parallel fashion. Assume that the path 1-2-4-8 is selected and the initial domains of the 1. Identify criteria's from begin to end nodes. $I1 <j1, j1> = k1$ then 2. Identify min and max values in the path allot to min and max variable. 3. Determine fixed values. “k1” fixed value obtained on node2 allotted to variable k1. 4. Determine fixed

**Fig. 1** Control flow graph**Table 2** Derived range of values for variable I1, J1, K1

Test table			
Variable I1	Variable J1	Variable K1	Test cases/path
0–30	50	10	T1/P1
0–9	10–50	10	T2/P2
10–30	0–30	20	T3/P3
30	0–50	20	T4/P4

values. “k1” fixed value obtained on node2 allotted to variable k1. 5. Then make use of obtained range to derive reduced test cases for all unique paths as given in Fig. 1 (Table 2).

5 Result Evaluation

Here, the proposed techniques is contrasted with the existing technique get split with respect to generated 1. total checklists or test cases. 2. Total depletion in test cases. 3. Comprehensive bugging time. 4. Fault detection rate by proposed prioritization technique.

Assume domain as $i1(0..30)$, $j1(0..50)$, $k1(0..40)$

where F1-> fault value entered less than minimum range value for $i1 = -1$ where F2-> fault value entered higher than maximum range value for $i1 = 31$ where F3-> fault value entered less than minimum range value for $j1 = -4$ where F4-> fault value entered is higher than maximum range value for $j1 = 51$ where F5-> fault value entered is less than minimum range value for $k1 = -5$ where F6-> fault value entered is higher than maximum range value for $k1 = 41$. Then for faults in Table 3, severity is given which is mentioned in Table 3. RFT = $Nj/TIMEj * 10$ is shown in

Table 3 No. of faults exposed and time required for each fault

Test cases/faults	T1	T2	T3	T4	NORDER	REVORD	PROP
F1				*	4	1	1
F2				*	4	1	1
F3			*	*	3	2	2
F4	*	*	*	*	1	4	3
F5			*		3	2	2
F6	*	*			1	4	3
No. of faults	2	2	3	4			
Time	2	3	3	4			
Severity	4	6	8	10	16	14	12

Table 4 Evaluated test case ranking value

Test cases	TCR = RFD + PFD + RDA
T1	17.33
T2	13.99
T3	23
T4	26.66

Table 5 Derived fault rate, PFD, RDA, and test case rank

Test cases	RFT	PFD	RDA	TCR
T1	10	3.33	4	17.33
T2	6.66	3.33	4	13.99
T3	10	5	8	23
T4	10	6.66	10	26.66

Table 5. PFD = NJ/total number of faults * 10 RDA = NJ * SJ/TJ TCR = RFD + PFD + RDA 2 is shown in Tables 4 and 5.

Code for Calculating Time Required to Execute Fault in Path1

```
import java.util.Date;
import java.util.concurrent.TimeUnit;
public class test{
public static void main(String []args){
int m1,m2,m3,t;
m2=51;
m3=10;
long startTime=System.currentTimeMillis();
if((m2==50) && (m3==10))
{for(m1=0;m1<30;m++)
{if(m2>=m1)
{ m3=10;
```

```

if(m2>=m3)
{t=m1+m2+m3+25;
System.out.println("m2 big"+t);}} } else{
System.out.println("error");
}long endTime=System.currentTimeMillis();
long timeElapsed = endTime - startTime;
System.out.println("Execution time in milliseconds: "+ timeElapsed);
}}

```

APFD = $1 - (TF_1 + TF_2 + \dots + TF_M)/M * N + 1/2 * N$. For no order, apfd = $1(4 + 4 + 3 + 1 + 3 + 1)/6 * 4 + 1/2 * 4 = 1 - 0.625 + 0.125 = 46\%$. For reverse order, apfd = $1(1 + 1 + 2 + 4 + 2 + 4)/6 * 4 + 1/2 * 4 = 1 - 0.3125 + 0.125 = 54\%$. For proposed order, apfd = $1(1 + 1 + 2 + 3 + 2 + 3)/6 * 4 + 1/2 * 4 = 1 - 0.4583 + 0.125 = 58\%$ as in NEXT. we parallelize test cases by referring test table given in Table 2. Now, variable i is used in three paths so range = total number of interval/3. Now, variable j is used in four paths so range = total number of interval/4. Since k is constant, we not divide k. So, range of I is spilt into three parts i1) 0...10 i2) 11...20 i3) 21...30, similarly range of j is spilt into four parts j1) 10...20 j2) 21...30 j3) 31...40 j4) 41...50. Now, when we execute them in a parallel fashion, we have total number of test cases = $[31 * 51 * 41] * 4 = 259,284$. Reduced test case for path1 = 31, but using existing technique test cases = 651, Reduced overall test cases = $[31 * 1 * 1 + 10 * 41 * 1 + 21 * 31 * 1 + 1 * 51 * 1] = [31 + 410 + 651 + 51] = 1143$, as shown in Table 2. Assigning each test case take constant 0.5 s, we observe that without parallelization execution time required for reduced test cases is $1143 * 0.5 = 571.5$. So, total number of test cases without test case reduction = 259,284 and execution time = 129,642. But for sequential execution Reduced test cases = 1143, and execution time = 571.5 and in case of PRIORITIZED parallel execution test cases = 1143 and execution. Time = 142.875 where processor having equal capacity and fault detection rate is more, and time required to detect all required FAULTS WILL BE considerably low because test cases are exposed in prioritized order as mentioned in Tables 6 and 7.

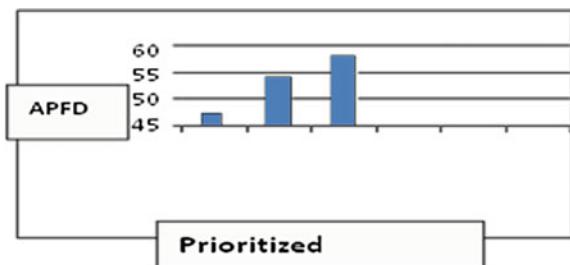
Table 6 Test case execution order of different ranking approaches

No order	Reverse order	Proposed order
T1	T4	T4
T2	T3	T3
T3	T2	T1
T4	T1	T2

Table 7 Calculated apfd values for different ranking approaches

Ranking techniques	apfd%
No priority	46
Reverse order	54
Proposed order	58

Fig. 2 Comparison between different prioritized techniques No, Reverse, and Proposed order



6 Conclusions

Each algorithm has its own significance as well as drawbacks ddr works on specific domain and split points, Ping-pong and other existing techniques results in more number of test cases, compilation, time, more effort and cost but proposed technique over performed by reducing number of test cases, prioritizing them and revealing more number of faults by assigning test case rankings, apfd calculation and then deriving prioritized reduced test cases as in Fig. 2, and finally by giving the test case rankings to achieve optimized performance and parallelized and prioritized test cases to reduce overall running time to execute and total budget required to perform testing.

7 Limitation

Our technique requires each and every path to execute serially to observe control flow and each and every path is to be examined to find out all possible constraints, and analyze relationships between variables, so it will take more time and memory to store the result, It is effective when the variables are there with constant and fixed values, and it works well for parallel execution where we can save memory and increase overall speed of execution.

References

1. Wang, R., Qu, B., Lu, Y.: Empirical study of the effects of different profiles on regression test case reduction. *IET Softw.* **9**(2), 29–38 (2015)
2. Biswas, S., Mall, R.: Regression test selection techniques: a survey. *Informatica* **35**(2), 289–321 (2011)
3. Boehm, B., Huang, L.: Value-based software engineering: a case study. *IEEE Comput.* **36**(1), 33–41 (2003)
4. Arnican, V.: Complexity of equivalence class and boundary value testing methods. *Int. J. Comput. Sci. Inf. Technol.* **751**(3), 80–101 (2009)

5. Abhijit, A., Sawant, Bari, P.H., Chawan, P.M.: Software testing techniques and strategies. *IJER* **2**(3), 980–986 (2012)
6. Jeng, B., Weyuker, E.J.: A simplified domain-testing strategy. *ACM Trans Softw. Eng. Methodol.* **3**(3), 254–270 (1994)
7. Gyimóthy, A.B., Forgács, I.: An efficient relevant slicing method for debugging. *SIGSOFT Softw. Eng. Notes* **24**(6), 303–321 (1999)
8. Saifan, A.A., Alsukhni, E., Ayat, H., Sbahi, A.L.: Test case reduction using data mining technique. *Int. J. Softw. Innov.* **4**, 56–70 (2016)
9. Assi, R.A., Masri, W., Zaraket, F.: A user-defined coverage criterion for test case intent verification. *Softw. Test. Verif. Reliab.* **26**, 460–491 (2016)
10. Kumar, A.: Evaluation of software testing techniques through software testability index. *AKGEC* **3**, 342–349 (2016)
11. Sharma, A., Patani, R., Aggarwal, A.: Software testing using genetic algorithms. *IJART* **7**, 1–13
12. Yadav, P., Kumar, A.: Software testing with different phases of SDLC. *IRTER* **2**, 118–122 (2016)

Hardware Realization of a Single-Phase-Modified P-Q Theory-Based Shunt Active Power Filter for Harmonic Compensation



**Krishna Veer Singh, Ravinder Kumar, Hari Om Bansal
and Dheerendra Singh**

Abstract This paper aims the harmonic compensation of inductive load using a single-phase Shunt Active Power Filter (SAPF). The instantaneous reactive power (p-q) theory is used in the controller to control the SAPF. This theory applies effectively for both, single- and three-phase systems. This theory is modified to achieve effective performance in single-phase systems. The system proposed has been simulated and analyzed in MATLAB, and then a lab prototype is developed and controlled using Arduino-based microcontroller. The response of the proposed system to variation in load is fast and takes care of sudden switching on/off.

Keywords Total harmonics distortion · Instantaneous reactive power theory · Microcontroller · Hysteresis current controller shunt active power filter

1 Introduction

Power quality has been one of the major problems in today's power industries. Since the fossil fuels are on the verge of extinction, there is a switch to renewable sources of power such as solar and wind energy, which makes the power quality ever so important. Most of the AC loads are inductive in nature and draw reactive power, which results in poor voltage quality. The issues of the power quality, include mainly sag and swell in voltage, unbalancing of voltage and current, voltage flicker, and

K. V. Singh (✉) · R. Kumar · H. O. Bansal · D. Singh

Department of Electrical and Electronics Engineering, Birla Institute of Technology and Science, Pilani, Rajasthan, India

e-mail: kriss.singh50@gmail.com

R. Kumar

e-mail: rksharmabits@gmail.com

H. O. Bansal

e-mail: hbansal@pilani.bits-pilani.ac.in

D. Singh

e-mail: dhs@pilani.bits-pilani.ac.in

harmonics. The SAPF used to mitigate the poor quality and enhance the reliability of sensitive loads [1–3].

SAPFs have various advantages over passive filters like small in size and don't need to be tuned for specific harmonic as they themselves adjust to the change in harmonic and compensate them. SAPF don't need much commissioning, and their attractive feature is to instantaneously compensate the harmonics and improve the power factor [4].

Various improvements have been accomplished in this field in converter design, switching techniques, compensation algorithms, etc. Many strategies have been put forward to control and mitigate the power quality issues for single-phase system. Table 1 shows the compassion of various control strategy.

From Table 1, it is clear that the instant action and reaction of instantaneous reactive power theory (the p-q theory) makes it suitable for SAPF control. Among various available controlling methods, p-q theory best suits to generate reference signals for single-phase SAPF system and works well in time domain. An SAPF based on p-q theory control is presented in [6–8] to evaluate its capability to mitigate the harmonics in the provided systems.

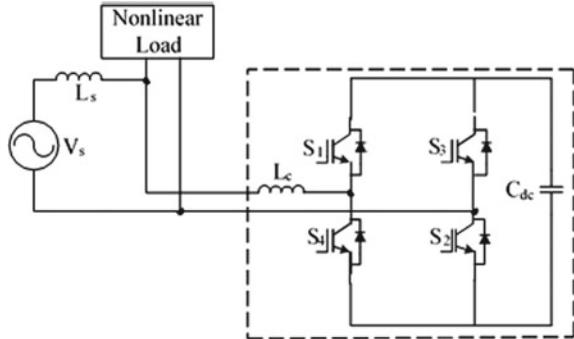
In this work, modified instantaneous theory-based SAPF is investigated for the applications, which are based single-phase system. A single-phase SAPF framework comprises of a 4-switch bridge structure along with a self-supporting DC bus inverter is appeared in Fig. 1.

The simulation has been carried out in MATLAB/Simulink environment. The applications of this system are in electric traction, low-power circuits, and rolling stocks. The aim of this paper is to contribute to the development of low-cost SAPF using Arduino-based microcontroller according to the specifications mentioned in the subsequent section. The simulations result proves that during distorted and unbalanced supply voltage and loading, SAPF is capable enough to mitigate the harmonics to a level as mentioned in power quality standards.

Table 1 Shows the compassion of various control strategy [5]

Condition	Instantaneous reactive power (p-q) theory	Synchronous reference frame	Sine subtraction	Fast Fourier transform
Steady-state response	Poor	Good	Excellent	Excellent
Transient response speed	Excellent	Good	Good	Excellent
Transient response quality	Good	Good	Poor	Poor
Requires voltage sensors	Yes	No	No	No
Number of filter stages	2	3	3	0

Fig. 1 Power circuit of single-phase SAPF



2 Single-Phase Active Power Filter

The system configuration of the proposed SAPF is shown in Fig. 1. SAPF comprises of a DC capacitor with a voltage-source inverter (VSI). The coupling inductor is used to connect the VSI at the Point of Common Coupling (PCC). SAPF might be viewed as a perfect current source associated at PCC. The reference current of the source that is required from supply subsequently after the compensation is minimum sinusoidal current, the microcontroller uses the signal of current and voltage with a specific end goal to create the reference signal of current. The current supplied after filtration is basically the difference of the reference signal of the current and load current. The hysteresis controller feeds both reference and actual current after filtration in a way so that it generates required signals for IGBT switches in the VSI.

3 Modeling of the Proposed System

The p-q theory is associated with the active and the reactive power analyzed in time domain. This theory is mostly used in three-phase systems, sometimes with or without a neutral wire. This theory gives a more accurate result for both steady state and transient condition [9, 10]. The source voltage of single phase in $\alpha - \beta$ coordinates with a $\frac{\pi}{2}$ lead can be represented as below

$$\begin{bmatrix} V_{s\alpha}(\omega t) \\ V_{s\beta}(\omega t) \end{bmatrix} = \begin{bmatrix} V_s(\omega t) \\ V_{s\beta}(\omega t + (\pi/2)) \end{bmatrix} = \begin{bmatrix} V_m \sin(\omega t) \\ V_m \cos(\omega t) \end{bmatrix} \quad (1)$$

The load current representation in $\alpha - \beta$ coordinates can be demonstrated in the form of equation as below

$$\begin{bmatrix} i_{L\alpha}(\omega t) \\ i_{L\beta}(\omega t) \end{bmatrix} = \begin{bmatrix} i_L(\omega t + \varphi_L) \\ i_L(\omega t + \varphi_L + (\pi/2)) \end{bmatrix} \quad (2)$$

The active and reactive powers can be represented as

$$\begin{bmatrix} p(\omega t) \\ q(\omega t) \end{bmatrix} = \begin{bmatrix} V_{s\alpha}(\omega t) & V_{s\beta}(\omega t) \\ -V_{s\beta}(\omega t) & V_{s\alpha}(\omega t) \end{bmatrix} \cdot \begin{bmatrix} i_{L\alpha}(\omega t) \\ i_{L\beta}(\omega t) \end{bmatrix} \quad (3)$$

The $p(\omega t)$ and $q(\omega t)$ can be expressed as

$$p(\omega t) = p_{-DC}(\omega t) + p_{-AC}(\omega t) \quad (4)$$

$$q(\omega t) = q_{-DC}(\omega t) + q_{-AC}(\omega t) \quad (5)$$

where $p_{-AC}(\omega t)$ and $q_{-AC}(\omega t)$ represents the AC components responsible for harmonic power and $p_{-DC}(\omega t)$ and $q_{-DC}(\omega t)$ represents the DC components responsible for instantaneous active and reactive power. By taking the inverse of (3), the reference compensating current can be calculated as

$$\begin{bmatrix} i_{c\alpha}(\omega t) \\ i_{c\beta}(\omega t) \end{bmatrix} = \begin{bmatrix} V_{s\alpha}(\omega t) & V_{s\beta}(\omega t) \\ -V_{s\beta}(\omega t) & V_{s\alpha}(\omega t) \end{bmatrix}^{-1} \cdot \begin{bmatrix} -p_{-AC}(\omega t) \\ -q(\omega t) \end{bmatrix} \quad (6)$$

$$\begin{bmatrix} i_{c\alpha}(\omega t) \\ i_{c\beta}(\omega t) \end{bmatrix} = \frac{1}{V_{s\alpha}^2(\omega t) + V_{s\beta}^2(\omega t)} \times \begin{bmatrix} V_{s\alpha}(\omega t) & -V_{s\beta}(\omega t) \\ V_{s\beta}(\omega t) & V_{s\alpha}(\omega t) \end{bmatrix}^{-1} \cdot \begin{bmatrix} -p_{-AC}(\omega t) \\ -q(\omega t) \end{bmatrix} \quad (7)$$

From Eq. (7), it is evident that the reference current is only dependent on the supply voltage profile. If the supply voltage has been distorted, then the extracted reference signal will also be improper. This is a major shortcoming of three-phase p-q theory.

In order to evaluate the error, DC link voltage is compared with reference voltage, and then it is minimized using a Proportional Integral (PI) controller. The references current are compared with supply current and error is achieved and this error is then passed the hysteresis current controller to provide the gate pulses for single-phase SAPF.

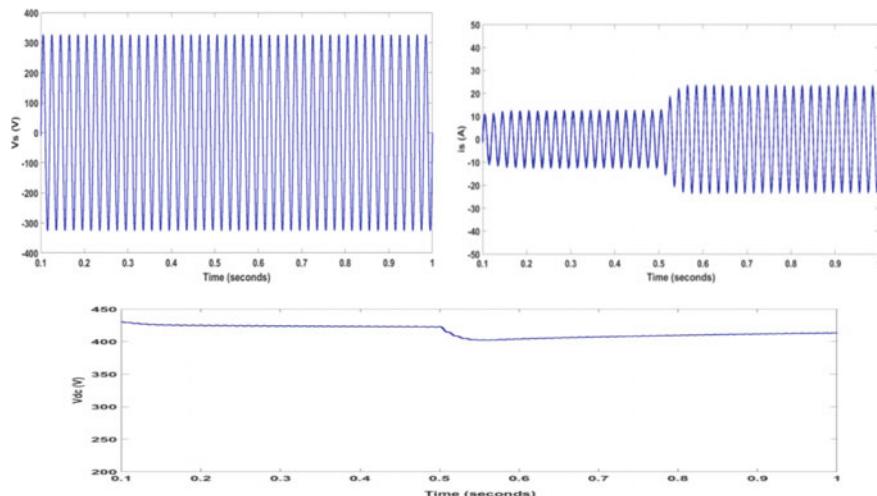
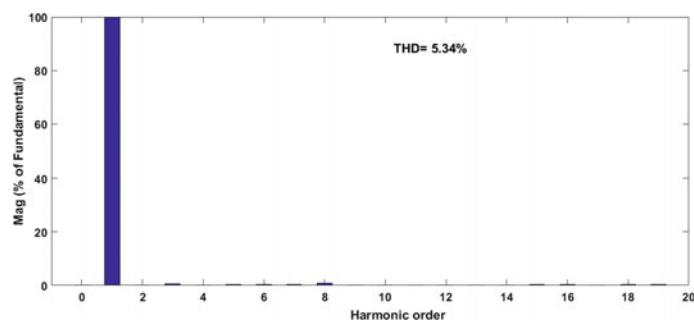
4 Results and Discussion

Simulation is carried in MATLAB/Simulink and hardware validation using Arduino-based microcontroller. Table 2 shows the system parameters.

Figure 2 shows the performance of SAPF under load-varying conditions in terms of source voltage (V_s), source current (i_s), and DC link voltage (V_{dc}). Figures 3 and 4 represent the harmonic spectrum of source and load currents, respectively. The algorithm's validity and feasibility has been tested in the laboratory via a prototype. Figure 5 provides the performance of the single-phase SAPF with the Arduino

Table 2 The system parameters

System	Parameters
Single-phase supply	RMS voltage = 230 V, Source resistance = 0.5Ω Source inductance = 0.1 mH , Supply frequency = 50 Hz
Nonlinear load components	Diode-bridge rectifier with $R_1 = 3 \Omega$, $L_1 = 3 \text{ mH}$, $C_1 = 1000 \mu\text{F}$, variable frequency drive with induction motor constant torque = 200 N/m^2 , moment of inertia $J = 0.005 \text{ kgm}^2$
Active power filter components	DC link capacitor voltage = 585 V

**Fig. 2** Shows the performance of SAPF under load-varying condition**Fig. 3** Harmonics spectrum of source current

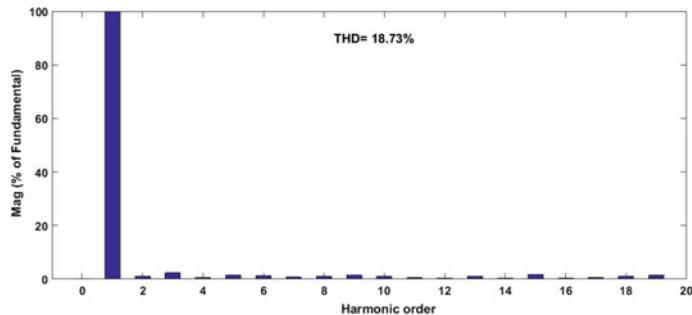


Fig. 4 Harmonic spectrum of load current

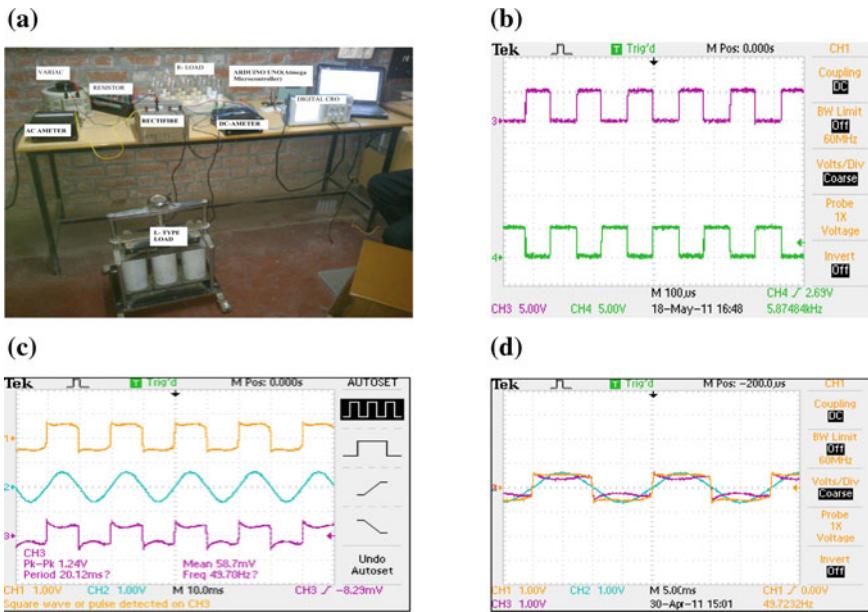


Fig. 5 **a** Hardware setup of single-phase SAPF **b** gate pulses **c** CH1: load current, CH2: reference source current, CH3: fundamental component of load current **d** load current, fundamental component of load current and source current

microcontroller in terms of load current, reference source current and fundamental component of load current. The advantage of this microcontroller is of its very low cost. The proposed control algorithm reduces THD and maintains the power factor near unity.

5 Conclusion

The designed and implemented single-phase SAPF minimizes THD and maintains the unity power factor in the grid. The line current after the compensation is sinusoidal, proving the best performance of control strategy used. The experimental results presented in this paper show that the developed SAPF suits well for application like residential areas, medicals, and offices, which are connected to the power grid by single-phase installations. Hence, on the basis of the experimental investigation and analysis, this manuscript proves the feasibility of single-phase modified p-q theory for practical applications. The approach reported in this paper can be used for industrial purposes due to its high reliability and cost-effectiveness.

References

1. Khadkikar, V., Singh, M., Chandra, A., Singh B.: Implementation of single-phase Synchronous D-Q reference frame controller for shunt active filter under distorted voltage condition. In: Joint International Conference on Power Electronics, Drives and Energy Systems, pp. 1–6 (2010)
2. Srinath, S., Prabakaran, S., Mohan, K., Selvan, M.P.: Implementation of Single Phase Shunt Active Filter for Low Voltage Distribution System, pp. 295–300 (2010)
3. Wu, J.C.: Utility-current feedforward-based control for a single-phase active power filter. *Int. J. Electron.* **98**(2), 185–196 (2011)
4. Ghodawat, S., Atigre, I.: Development of single phase shunt active power filter. In: International Conference on Inventive Communication and Computational Technologies, pp. 351–355 (2017)
5. Kumar, R., Bansal, H.O.: Shunt active power filter: current status of control techniques and its integration to renewable energy sources. *Sustain. Cities Soc.* **42**, 574–592 (2018)
6. Ranjbar, M., Jalilian, A.: Modified single-phase instantaneous reactive power theory: formulation, simulation, and experimental verification. In: Power Quality Conference (PQC), no. 3 (2010)
7. Ranjbar, M., Jalilian, A.: Implementation of a single-phase shunt active power filter under nonsinusoidal voltage source. In: Proceedings—2010 18th Iranian Conference on Electrical Engineering ICEE, no. 2, pp. 759–764 (2010)
8. Altus, J., Michalik, J., Dobrucky, B., Viet, L.H.: Single-phase power active filter using instantaneous reactive power theory-theoretical and practical approach. *Electr. Power Qual. Util.* **XI**(1), 33–38 (2005)
9. Kunjumuhamed, L.P., Mishra, M.K.: Comparison of single phase shunt active power filter algorithms. In: IEEE Conference, pp. 1–8 (2006)
10. Lin, B.R., Huang, C.L.: Generalised single-phase p-q theory for active power filtering: simulation and DSP-based experimental investigation. *IET Power Electron.* **2**(1), 67–78 (2009)

Development and Gesture Control of Virtual Robotic Arm



Vineeth Veeramalla, Jagriti Mishra, C. S. Meera, Varnita Verma and Mukul Kumar Gupta

Abstract In this study, we propose the development and control of a gesture-controlled four degrees of freedom (4 DOF) virtual robotic arm. The virtual robotic arm mimics the motion of human arm with the help of physical sensors like gyroscope and flex sensors attached to the wrist and elbow of the human arm. The position and orientation data from the sensors are converted to corresponding joint angles and are transmitted serially to bring the same motion in the virtual robotic arm. The position control for the virtual arm is achieved using an PID controller developed in MATLAB using Arduino as an interface. The controller gains are tuned to reduce the errors and optimize the system performance.

Keywords Robotics · Gesture control · Sensors · PID control

1 Introduction

Robotic technologies have played and will continue playing important roles in helping to solve real-life problems. In which, Human–Machine Interaction is one of the most interesting fields in the development of robots. The method put forward, here, is a small approach to interaction between a human and a robotic system. The proposed technology gives the human–machine interaction in which a robot arm mimics the movements of a human arm accurately as much as possible [1, 2].

Gesture-based human–robot interaction is not a new technology, but the strategy for its control is different compared to earlier researches. Many studies are being done to control this system using rotary encoders for each joint variable, image processing and several other methods. Most of the cases used accelerometer and gyroscope [3]. Studies are focusing on various simulator-based platforms as they can bring down the cost of testing various controllers and requirements on the real system. Once

V. Veeramalla · J. Mishra · C. S. Meera (✉) · V. Verma · M. K. Gupta

Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies (UPES), Dehradun, India
e-mail: meerasunil007@gmail.com

tested successfully, these methods proposed can be implemented on real robotic arms. Furthermore, in teleoperation and control of robotic systems, operator sitting at a remote terminal sees the robotic simulator and performs the operation.

There are several ways of controlling a robotic arm. In the last few years, so many researches have been done on controlling of robotic arm through various methods. The applications developed for robot control are server-based and therefore, they can be remotely accessed and controlled from any location in the world. In general, according to the task space function to be performed, the controller takes the input from the operator and moves the robotic arm to the desired position. The main difficulty in such method is to calculate the input values for motors to move to a particular position or to achieve a particular movement. This is easily achieved by our project. In this project, we are trying to make a simple human hand mimicking robotic arm using 3 DOF gyroscope and a flex sensor. This type of robot does not require any kind of training as the robotic arm mimics hand movement of the operator simultaneously. Here, the interfacing is carried out through wired communication but can be changed to wireless easily [4, 5]. Accelerometer equipped with a gyroscope will make movement smoother [6].

Here, the end-effector motion control has been considered. However, the end-effector motion can also be controlled by incorporating the required number of sensors and can be designed for any suitable application such as gripping.

2 System Description

A robotic arm is a mechanical structure, which resembles a human arm and is used to contact a material or object without direct contact by the human. Such robotic arms are typically used in industrial applications such as painting, welding, assembly of parts, etc. Whereas, human hand resembling robotic arm, on the other hand, are mostly used in fields, working with radioactive material where direct human contact is dangerous and in the medical field for surgical robots where human hand-like movements are required. The human arm has 7 degrees of freedom (DOF), 3 degrees of freedom at shoulder, 1 degree of freedom at elbow and 3 degrees of freedom at wrist joint [7]. But here, the proposed work only consists of only a 4 DOF of robotic arm which can perform 2 degrees of freedom at shoulder, 1 degree of freedom at elbow and 1 degree of freedom wrist joint movements. The sensors used for detection of position and orientation of human operator are as follows.

Gyroscope (MPU6050) helps to determine the orientation: A gyroscope measures the force of the gravity and determines the orientation. Accelerometer measures the velocity based on the principle of conservation of momentum. Whenever the 3-axis accelerometer is used in combination with a 3-axis gyroscope, the output is smooth and responsive.

Flex sensor is a type variable resistor. The resistance of flex sensor changes when it is bent. The bend radius is proportional to the resistance of the flex sensor. Carbon

resistive elements with thin flexible substrate are used inside the flex sensor. More carbon means less resistance. Thus, smaller the radius higher the resistance value. Flex sensor basically works in two directions. When the sensor is bent in compressed direction, the resistance across it decreases and the resistance of the flux sensor increases as it is bent in tensile direction [8, 9].

3 Gesture Control

Figure 1 represents the block diagram representation of the gesture control system. In this, the values are taken from gyroscope and flex sensor placed on a human arm in such a way that the values will be converted into the angles of servo motors, which is fixed in the robotic arm and will exactly mimic the human arm. This approach is not new but we are trying to reduce the wearability of human operator and get the values as accurate as possible with our project [10, 11].

The main target is to reduce the wearability of human controller. In this project, it's a small model with a gyroscope and flex sensor, which will reduce the wearability and tries to give as accurate as possible results and the delay provided by it as less as possible.

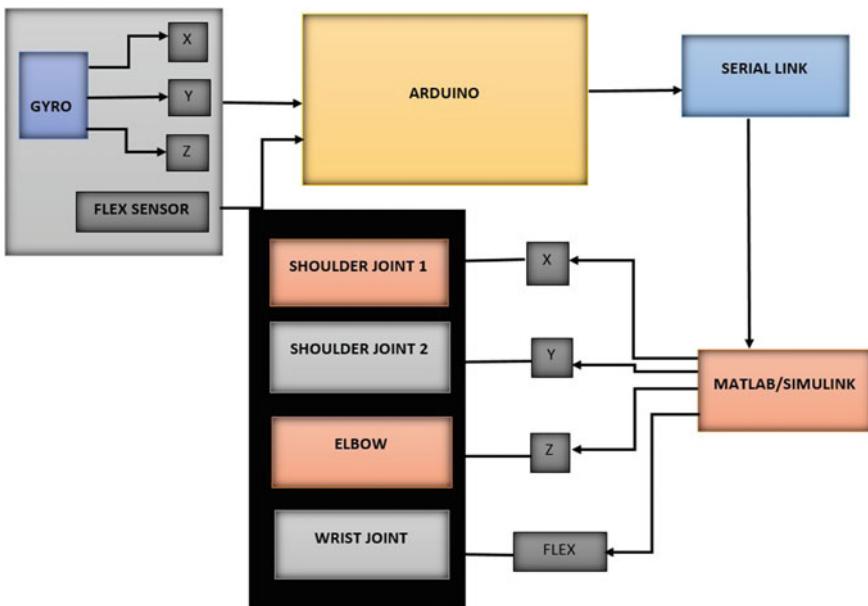


Fig. 1 Block diagram representation of gesture control system

3.1 Positioning of Sensors

From the experiment, three values has been taken from gyroscope X, Y, Z, and for proper results, the positioning of sensors is done further as gyroscope will be placed on the end effector in between the wrist and knuckles, and the flex sensor is placed exactly at the elbow joint to get the exact values from the human arm to control the robotic arm.

An accelerometer measures the inertial force and a gyroscope sensor measure the orientation based on gravity. The MPU6050 is equipped with both (a) 3-axis accelerometer and (b) a 3-axis gyroscope sensor at the same time in one breakout board. This sensor comes under MEMS (Microelectromechanical Systems) technology and I2C-bus is used for interface it with the Arduino board. The sensor detects in X, Y and Z axis and the output ranges from -17000 to $+17000$. And further, these values are converted to angle from 0 to 180 of a servo motor. Similarly, the flex sensor values are also converted to servo angles from 0 to 180° . Figure 2 shows the configuration of the virtual robotic arm and position-control system.

Converting Gyroscope Values to Servo Angles: Gyroscope gives output from -17000 to $+17000$. The difference divided by number of degrees gives the value for each degree.

$$\begin{aligned} & (17000 - (-17000))/180 \\ & = 34000/180 \\ & 1 \text{ degree} = 188.88 \end{aligned}$$

For each degree of inclination of the gyroscope gives 188.88 as output.

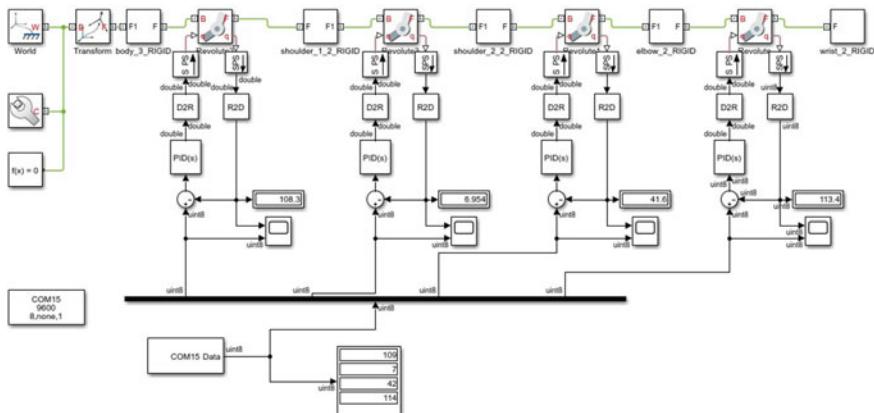
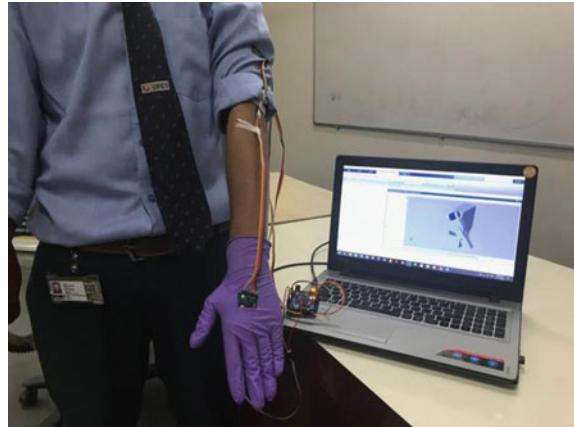


Fig. 2 Configuration of the virtual robotic arm and position-control system

Fig. 3 The system demonstration showing the gesture mimicking of human arm by the virtual arm



Converting Flex Sensor Values to Angles: The output values of the flex sensor depend on the type of resistor used. The difference between the maximum value and minimum value divided by no. of degrees gives value for each degree.

4 Experimental Results

The gyroscope and flex sensor data were transmitted serially to obtain position control of the virtual robotic arm. The sensor was initially calibrated to obtain the joint angles from 0 to 180 (degrees). Figure 3 shows the demonstration of gesture mimicking by the virtual robotic arm.

The data obtained was filtered and given to a PID controller. Figure 2 shows the mimicking of gestures between human arm and the virtual robotic arm. The controller gains were tuned to achieve position tracking in the MATLAB/SIMMECHANICS platform. The output results showed good tracking of the input sensor values. Figure 4 shows the position tracking of the four joints of the virtual robotic arm. The shoulder joint 1 and 2 and wrist joint were controlled through the gyroscope values and elbow joint were controlled through flex sensor values.

5 Conclusion and Scope for Future

The proposed project is an approach to control a robotic arm using flex sensor and a gyroscope, which is more instinctual and easier to work with, besides offering the possibility for the robot to mimic the human operator. Using this system any inexperienced operator can control robotic arm quickly and easily. Also, it can be used in applications where precise control and mimicking of human arm motion is

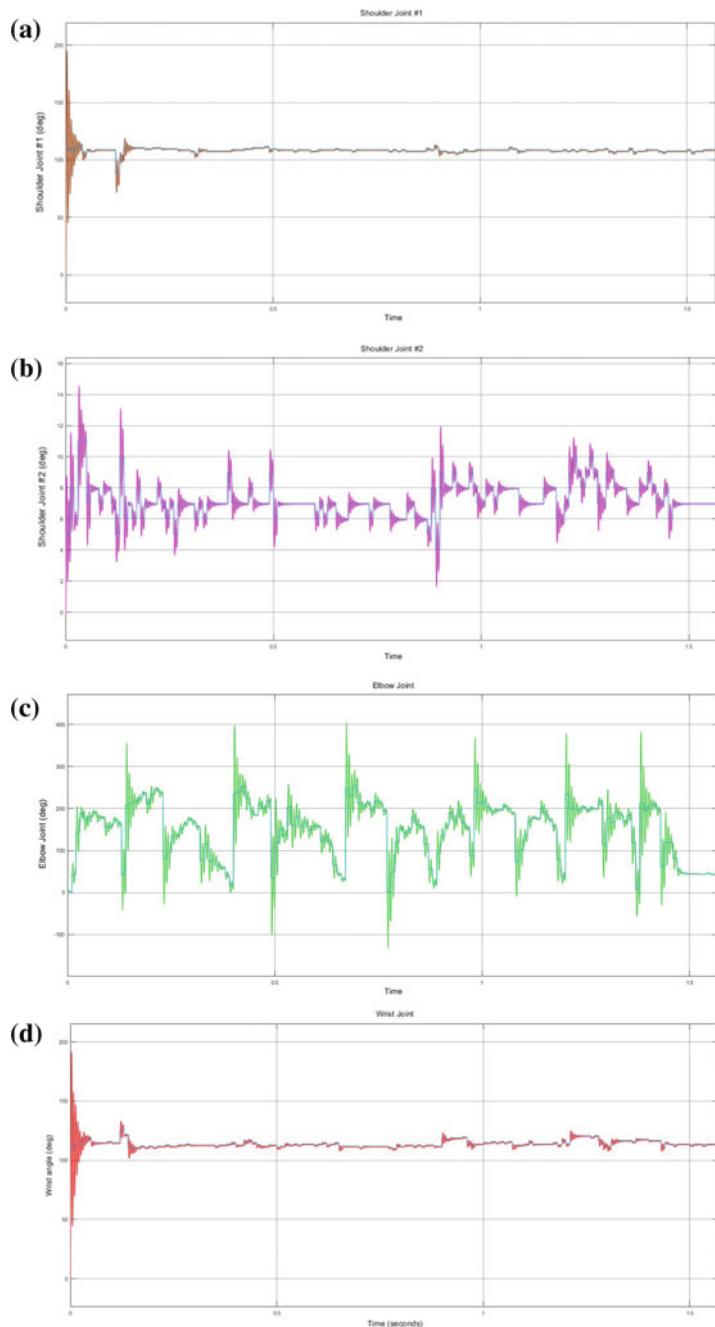


Fig. 4 **a** Position tracking of shoulder joint 1. **b** Position tracking of shoulder joint 2. **c** Position tracking of elbow joint. **d** Position tracking of wrist joint

required, this method can be easily implemented. It provides a more flexible control mechanism. Although, the mimicking of human arm is achieved but the problem of noise and jerks can be there, which can be reduced by calibration of sensors and using a much precise smoothing algorithm. This process reduces the material cost, human labour and power consumption. Further research can be done for more accurate results. The physical model of such a system can be made bilateral and haptic feedback, which can be achieved using force feedback sensors. This human arm mimicking technology is not only restricted for industrial purposes but also can be applicable in the field of rescue and defence operations. Further improvements can be made and study for prosthetics is also possible. More robust systems can be made for underwater applications, integrating with wireless technology and camera feedback systems.

Acknowledgements The authors would like to thank Robotics Centre of Excellence, UPES, Dehradun for the continuous support and guidance.

References

1. Brahmani, K., Roy, K.S., Ali, M.: Arm 7 Based Robotic Arm Control By Electronic Gesture Recognition, vol. 4, no. April, pp. 1245–1248 (2013)
2. Vyawahare, V.M.: Design of a prosthetic arm using flex. *Int. J. Electron. Commun. Eng. Technol.* **8**(2), 1–6 (2017)
3. Sekhar, R., Musalay, R.K., Krishnamurthy, Y., Shreenivas, B.: Inertial sensor based wireless control of a robotic arm. In: 2012 IEEE International Conference on Emerging Signal Processing Applications, pp. 87–90 (2012)
4. Doshi, M.A., Parekh, S.J., Bhowmick, M.: Wireless Robotic Arm Using Flex Sensors, vol. 6, no. 3, pp. 1471–1474 (2015)
5. Maruthisagar, N.V.: MEMS Based Gesture Controlled Robot Using Wireless Communication, vol. 14, no. 4, pp. 185–188 (2014)
6. Bhuyan, A.I., Mallick, T.C.: Gyro-accelerometer based control of a robotic Arm using AVR microcontroller. In: 2014 9th International Forum on Strategic Technology, pp. 409–413 (2014)
7. Filiatrault, S., Cretu, A.: Human Arm Motion Imitation by a Humanoid Robot (2014)
8. Gautami, J., Sampada, A., Sahil, M., Dadhich, P.: Flex Sensor Based Robotic Arm Controller: Development, vol. 328, pp. 494–498
9. Hande, J.Y., Malusare, N., Harshaldarbhe, S.: Design for Robotic Hand Using Flex-Sensor, vol. 4, no. 12, pp. 2846–2850 (2015)
10. Malegam, K.D., D'Silva, M.S.: Mimicking robotic hand-arm. In: Proceedings—2011 Annual IEEE India Conference on Engineering sustainable solutions INDICON-2011 (2011)
11. Lovasz, E.C., et al.: Design and control solutions for haptic elbow exoskeleton module used in space telerobotics. *Mech. Mach. Theory* **107**, 384–398 (2017)

Real Virtual IDentification (RVID): Providing a Virtual, Secure and Anonymous ID Service to Indian Users



Amit Kumar Tyagi, G. Rekha and N. Sreenath

Abstract Today using (sharing) Aadhaar with several welfare schemes/organisations, provide several essential benefits (to get government schemes benefits) to Indian citizens. But on the other hand, revealing/sharing their Aadhaar number with government and organisation raised several issues like security and revealing of personal information of the user. For that the government makes several enhancements from time to time like recently, government forces and give an option to use virtual IDs irrespective of using Aadhaar number. But virtual ID is not fixed and a user can generate multiple VIDs, it means it is a confusing process to user/people belonging to rural area. So to solve this issue of VID, this work describes a novel approach to provide a secure and anonymous (yet transparent and immutable) identification management system, which provide a Real Virtual ID (RVID) with respect to every user's Aadhaar number. This ID (16-digit ID, i.e. containing combination characters and letters) can be used to take several benefits or can be used irrespective of sharing Aadhaar number with organisations. A prototype shows that such ID management system is feasible, and solves the problem of duplication of data. It is also immune to many ID attacks like guessing attack, etc.

Keywords Aadhaar · Social/welfare schemes · Governments · Attack · Identification provider · Anonymous · Transparent · De-duplication

A. K. Tyagi (✉)

School of Computer Science & Engineering, Lingayas Vidyapeeth, Faridabad 121002, Haryana, India

e-mail: amitkrtyagi025@gmail.com

G. Rekha

Department of Computer Science and Engineering, K L University, Hyderabad 500075, India
e-mail: gillala.rekha@klh.edu.in

N. Sreenath

Department of Computer Science and Engineering, Pondicherry Engineering College, Puducherry 605014, India
e-mail: nsreenath@pec.edu

1 Introduction

Till today, more than 100+ crore Aadhaar card has been issued by Government of India (GoI) to its citizens free of cost (for first time/creation time). Aadhaar has saved millions of rupees via reducing fraud in government offices or providing benefits/needs to respective users (based on this unique Aadhaar number). Every citizen (of India) contains this unique number as nationally, not globally. Basically, the actual UIDAI-Aadhaar number is 11 digits and not 12 digits. This doubt is clear here by ‘the first 11 digits of the 12 digit Aadhaar number (just a random generated number) displayed on an Aadhaar card is the actual UID Number and the 12th digit is the checksum associated with Verhoeff Algorithm scheme’ [1]. Note that in generation of an Aadhaar number, it does not contain any formula like other identities (like PAN, Passport, etc.), i.e. collection of pin code, demographic location or name, etc. Today in Aadhaar, Identification management, authentication, authorisations and service (IDs in short) are important research problems to work.

As discussed, Aadhaar does not contain any hologram, it just contains QR code which can be easily tempered. This tempered Aadhaar can be used in any illegal activities by any terrorists. Basically, Aadhaar was initially proposed and recommended by the Kargil Review Committee (KRC) to provide a unique identification number to its citizen or to every people. Today, Aadhaar is just used as ‘proof of identity’, not as a ‘proof of address’. Moreover, several issues and challenges have been raised with Aadhaar time to time by several politicians and social activists. For that government has taken several steps to improve people privacy. The government has provided several new features like locking your biometrics online, sharing of Virtual Identification (VID) to its citizens with respect to Aadhaar, i.e. to protect their personal information from strangers. In general, the VID is a temporary, revocable 16-digit random number which can be generated multiple times by users and have no expiry time. The VID will allow the user to authenticate transactions and e-Know-Your-Customer (KYC) services instead of providing their 12-digit Aadhaar number to organisations [2]. Note that, a generated VID works with indexing/mapping process with the respective Aadhaar number. But generating these ID or using Internet technologies is challenging to most of citizens, i.e. to rural people.

In authorising or verifying exact/genuine user, user’s registered mobile number and one-time password pair is still a very effective approach to do authentication and authorisation [2]. Unique Identification Authority of India (UIDAI) does not store any combination of password or usernames (of their citizens) in their database because most of the users use simple surname and password pair, which can be easily guessed by attackers. So, it (UIDAI) always works with One-Time Password (OTP) concept to verify the authenticity of user. In recent, providing of VID and sharing only this ID has been emerged as a revolution to citizens of India. But sharing of this VID does not provide enough security and privacy to its citizens. Hence in this work, we build a novel type of ID management and ID service, i.e. Real Virtual Identification (Real Virtual ID) with secure decentralised anonymity, unlinkability, transparency and immutability. Such IDs are necessary to provide a unique identification to peo-

ple with protecting their privacy. The proposed Real Virtual IDentification (RVID) is used for building a secure and anonymous yet transparent and immutable identification IDs via providing ID management to users/Aadhaar system, i.e. authentication, authorisation, storage, and ID service. In addition, the Real Virtual ID use one extra layer of security to reduce accessing time or updating time (to a database) of generated Virtual ID to users. Using this service, a user can be anonymous or can access all services without revealing his identity inside a crowd. The Real Virtual ID can also use characteristics to identify devices and products. Using Real Virtual ID, users can access several benefits, i.e. without revealing their Aadhaar number. Some of the benefits (of using Aadhaar) are: Aadhaar-based Direct Benefit Transfer (for receiving LPG Subsidy) for its citizens, Jan Dhan Yojana (opening of bank account at zero balance for poor people), Issuing of Passport in a few days (via reducing verification process), Digital Locker (a platform to keep your records online securely), Voter Card Linking (reducing fraud voters), Monthly Pension (identification of respective and correct people eligible for pension like disable, old-age, etc.) Provident Fund, Opening new bank account (for all people), Digital Life Certificate, Transfer of scholarship/fellowship amount to right candidate, SEBI, etc. [9]. Also, these benefits of Aadhaar number cannot be ignored now onwards as it has now been made mandatory for the most important day to day activities we do. For example, for tax/fund-related activities. Note that issuing a new mobile number/linking of mobile number now verdict given by Honourable Supreme Court (SC) is that organisations/government cannot ask for Aadhaar for linking (with their mobile number) from users, but it makes mandatory for PAN linking, i.e. for tax purpose, linking of driving license, for investments, for investments, for existing bank account holders, for making a financial transactions above Rs. 50,000 and so on.

Hence, all the above benefits show that why it is must for every Indian moving forward to use Aadhaar (12-digit unique number). The organisation of this paper is as follows: Sect. 2 discusses Virtual ID fundamentals, mandatory requirement for issuing Aadhaar and feature of virtual ID. Further, Sect. 3 discusses our purposed concept/approach with respect to Aadhaar, various issues with Aadhaar, etc., in detail. Then in Sect. 4, we provide an open discussion with respect to our proposed approach, i.e. Real Virtual IDentification. Finally, this work is concluded in brief in Sect. 5 with future possible enhancements.

2 Virtual ID Fundamentals

We provide the UIDAI virtual ID instead of our Aadhaar number to organisations/agencies and protect our basic Aadhaar details from outside world/malicious users/being accessed by someone else. Aadhaar Virtual ID is a 16-digit temporary code that can be used for Aadhaar authentications.

But a big question is raised here, i.e. why we need it (VID) here? In the past, there have been a number of cases (several times) where information from UIDAI/Aadhaar has been leaked. So, people became very much concerned about their information

stored with UIDAI. UIDAI has addressed the concerns of people and has come up with the virtual ID. Virtual ID (VID) is an alternative to Aadhaar number, provided by the Government of India to its citizens to protect their privacy/personal information against several attacks. When users/citizens use their respective virtual IDs instead of Aadhaar number, organisations/local agencies cannot collect the Aadhaar number of the respective user/applicant. All organisations/local agencies do authentication with virtual ID only. This virtual ID works based on mapping process, i.e. provides a relationship between VID and Aadhaar number [3]. VID provides anonymity to local agencies/organisations, i.e. preserve private information of user from the outside world. Using this VID, the Aadhaar details are not accessed by the outside world, i.e. VID is keeping the Aadhaar number and other details safe from being hacked. Aadhaar holders can generate VID either once or multiple times (expiration period is not fixed for VID) and use it (VID) to authenticate user for providing a service to respective user/Aadhaar holder. Once the authentication is complete, the user can regenerate his virtual ID (in future) so that the saved details with the organisations became useless. Hence, Fig. 1 shows connection among Central Identities Data Repository (CIDR), and Authentication User Agency (AUA) or KUA. Creating several VIDs put a burden to Aadhaar database and confusion with organisations, which is a serious issue. These Authentication User Agencies (AUAs) collect and generate Aadhaar number to its people/citizens after collecting their personal information and biometrics in their system [2]. Note that according to the Aadhaar Act 2016, a requesting entity means ‘an agency or a person that submits Aadhaar number and demographic information or biometric information, of an individual to the Central Identities Data Repository (CIDR) for authentication’ [4, 5]. Here, Authentication User Agency (AUA) is ‘an entity who is engaged in providing Aadhaar Enabled Services to Aadhaar holder (using authentication facilitated by the Authentication Service Agency (ASA))’ [5]. Hence, Fig. 1 shows that a requesting entity (such as AUA, KUA) connects to the CIDR through an Authentication Service Agency (ASA) (either by becoming ASA on its own/by contracting services of an existing ASA).

2.1 *Mandatory Security Requirements*

Several mandatory security requirements have been issued (or used) for a requesting entity/issuing entity to generate Aadhaar (in last years). Some of them are listed as:

- An Aadhaar number should be never used openly (as we have discussed in [9] that sharing an Aadhaar number is just like sharing a mobile number, but when we share openly our Aadhaar number with our physical card, then we lose our identity and demographic location also). And in the case of Aadhaar-enabled centre (or operators assisted devices), operators need to be authenticated via password, Aadhaar authentication, etc., mechanisms [6].
- Personal Identity Data (PID) blocks (collected for Aadhaar authentication) need to be encrypted during collection/capture without sending over a network. Also,

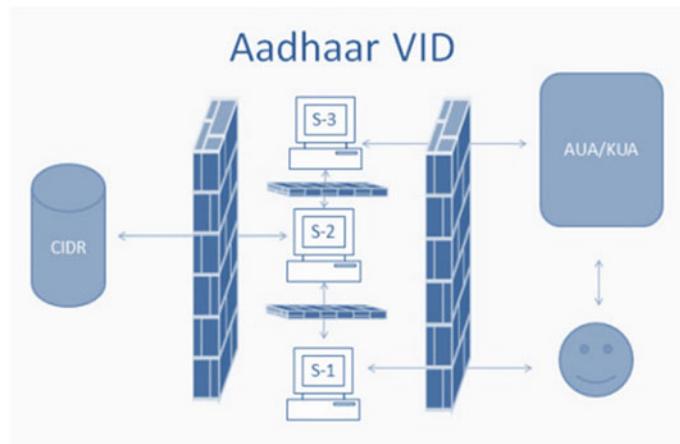


Fig. 1 Process of existed virtual identification mechanism

do not store the encrypted PID blocks (with updated/good key of combinations), unless it is for buffered authentication for a short span of time (currently this time is 24 h).

- Biometric and One-Time Password (OTP) data collected for Aadhaar authentication should not be stored on any permanent storage/database [5, 6].
- The metadata and the responses should be logged for audit purposes for detecting bogus users/reduce fraud users.

Note that network between AUA and ASA should be secured all the time. Here as discussed above, an AUA is ‘any entity that uses Aadhaar authentication to enable its services and connects to the CIDR through an ASA, whereas ASAs are entities that have secure leased line connectivity with the CIDR’ [5]. ASAs (ASAs work with a formal contract with UIDAI) transmit authentication requests to CIDR on behalf of one or more AUAs.

2.2 Features of the Aadhaar Virtual ID

Unique Identification Authority of India (UIDAI) has made various upgrades in its system through the VID. Some of the salient features of the virtual ID are:

- (a) The virtual ID is a temporary 16-digit code that will replace Aadhaar for authentication,
- (b) There is only one virtual ID issued at a time. When a new VID is issued, the old one is replaced out (or removed from database/record),
- (c) Agencies have to update their system to include VIDs by 1 June 2018,
- (d) Aadhaar number cannot be retrieved from the virtual ID,

- (e) There is no cap on the generation of virtual IDs,
- (f) It is not compulsory to generate VIDs. A person can furnish his Aadhaar instead of the VID,
- (g) Agencies cannot force applicants to provide Aadhaar number for e-KYC or verification,
- (h) Agencies have to take consent from the user for authentication using VID,
- (i) No agency is authorised to store the virtual ID or any other Aadhaar details taken for authentication,
- (j) The virtual ID is valid till the user generates a new one,
- (k) When you retrieve your Aadhaar, the last generated VID is sent to the registered mobile number.

Note that we can generate virtual IDs for other Aadhaar numbers as well for your family. A user can use VID from 1 June 2018 with organisations irrespective of sharing his Aadhaar 12-digit number [2]. The generated VID is completely temporary, revocable and 16-digit random number. It can be created by users multiple times. So storing different VIDs every time will create more confusion among local authorities. Also, it will put a burden/require additional database to store multiple VIDs, which may create a reason to lose user's privacy. Today, losing of privacy in computer technology/during making a communication process is a critical issue [8]. Hence, this section discusses Virtual ID system, which is provided by UIDAI and also explains several features of VID. Now, next section discusses our proposed approach or enhancement of VID in detail.

3 Real Virtual ID: Our Proposed System

Multiple Virtual Identifications put a burden on Aadhaar database or agencies/local organisations. To reduce this load and confusion and increasing efficiency and improving accessing cost, we proposed a novel approach, i.e. proposing real virtual identification using a token. Figure 2 demonstrates the work flow of our proposed approach. The following Fig. 2 describes the detail steps of our proposed approach of real virtual identification. Using pseudonyms or Providing extra layer as unique token reduce searching process in our Aadhaar database. When a user creates multiple VIDs for himself/his Aadhaar number, then it is difficult to match or indexing his generated VIDs to his respective Aadhaar number. It requires a lot of time and a complex process. In this case, a lot of irrelevant information is created for that respective user. So to reduce the complexity and improving the accessing or searching level in Aadhaar, we use an extra level of term, like providing a unique token to each user, i.e. based on user's Aadhaar number (see Fig. 2).

With providing this, we provide anonymity to users and we do not have to think about storage because this token has all possible match or generated virtual ID for every user.

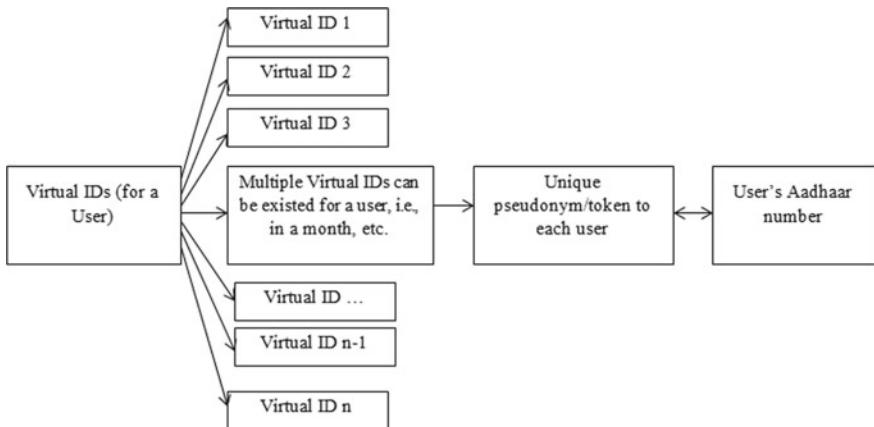


Fig. 2 Real virtual identification system

3.1 Virtual ID Generations

Generating Real Virtual ID (RVID) system is not a complex task. Real Virtual ID can be generated by any user/citizen of India any time from Intranet, Internet or from UIDAI website. Note that for generating a real virtual ID, a user should hold an Aadhaar card and he should also register a number with Aadhaar database. Then only, he/she can generate a virtual ID for himself/herself. Hence, there are some steps to generate Real Virtual ID are:

- Step 1: A user visit UIDAI's website at <http://uidai.gov.in/>.
- Step 2: User click on the 'RealVirtual ID (RVID) Generator' from Aadhaar services section.
- Step 3: User will be moved to a new VID Generation page <https://goo.gl/vFQgic>.
- Step 4: User need to enter his 12-digit Aadhaar Number and the security code.
- Step 5: Now user need to click on the 'Send OTP' button.
- Step 6: A One-Time Password (OTP) will be sent to user's mobile number (which is registered with UIDAI/Aadhaar database).
- Step 7: user enter the OTP and select the option to either 'Generate RVID' or 'Retrieve RVID'.
- Step 8: Now user click on the Submit button.
- Step 9: User will receive a message like 'Congratulations! Your VID Number Successfully Generated and sent to your registered mobile'.
- Step 10: Hence as the last step, user will get the message on his registered mobile number mentioning the 16-digit real virtual ID for Aadhaar number and the last 4 digits of Aadhaar is xxxxxxxx9870, generated at 25-09-2018:01:01:01.

Hence as output, we get a RVID like that (after Step 10), congratulations! Your RVID has been created which is 1024wesvb67er17s for your Aadhaar number xxxxxxxx9870.

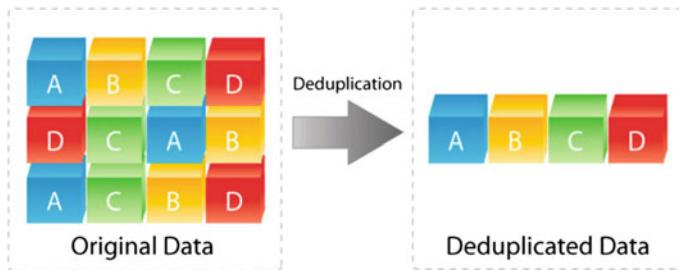


Fig. 3 De-duplication of data in generating RVID

3.2 Benefits of Real Virtual ID

We provide enough security and fast accessing and updating process/Aadhaar database. Also in our proposed system, several VIDs can be generated by a user but here, a unique token contains 16-digits, i.e. a combination of characters and numbers. The process to generate RVID is similar like generating simple VID, but here our system will allocate a unique token to each user, then it will generate VID to user based on this generated token.

De-duplication of data cannot be done by any agencies in our proposed system. Our approach eliminate of duplicate or redundant information to make accessing fast in Aadhaar database (see Fig. 3).

3.3 Issues with Real Virtual ID

As discussed above, RVID is an extra layer of security to protect Aadhaar card information or user's information. But, what about that entire information/Aadhaar card's information which has already been revealed or already has been shared with several local organisations in past? By the government, there is no expiry period is defined for a VID, i.e. multiple VIDs have been created by users with respect to their Aadhaar numbers. A VID (a 16-digit digital ID) is valid till next one VID is not generated. Generating and accessing of RVID depend on user's mobile number, i.e. to generate this ID, user receives an OTP to authenticate himself.

Hence, this section discusses our proposed approach with several benefits and issues raised in brief. Now, the next section will provide an open discussion with research communities with respect to previous existing ID, i.e. virtual and with our proposed approach, i.e. Real Virtual ID.

4 Open Discussion

Biometrics is the most useful discipline among all (except mathematics), used related to human characteristics for identification, authentication and access control. Using biometrics like fingerprint, palm veins, face recognition, DNA, palm print, hand geometry, iris recognition, retina, etc., considered for providing security to a system (Aadhaar database). Also, some other metrics are used to authenticate purpose (or for accessing a system) like voice, typing rhythm [7]. But among all of them, biometrics provides good and better accuracy (with technology advances). For authentication, we need to complete the authentication process in less time (in milliseconds), i.e. the time to get information is needed to be faster with affordable (low) price. Also, one essential feature of biometric authentication is needed to be detecting fake thumb impression/biometrics as quickly (for real time authentication). In several applications, biometrics is used to differentiate people and robots. But here, the main challenge is to detect non-human biometrics efficiently in less time with low cost. Differentiating human from robots is essential to use in applications like which requires anonymity but identifiable human natures. We will use methods like Principal Component Analysis (PCA), Wavelets and Correlations to test such Classification/Proposed System. Also as discussed above, in the proposed Real Virtual ID (RVID) system, captured biometrics is being used but not stored (more than 1 week). So, it becomes hard to get original biometrics after a particular time span. In addition, it also needs to reduce the probability of false positive in biometric collection. Note that, it is clearly impossible to have similar biometrics for two different persons/twins (within certain thresholds).

We have illustrated a novel approach of ID management and authorisation service using an extra token. We noticed that providing extra token/a unique token number is a good step (fit) for ID management and service. By using extra token, the updated system became immune to many attacks like background attack, homogeneity attack, an ID attacks, etc. Hence, a prototype implementation will demonstrate (in extension of this work) that RealVirtualID serves and overcome all Aadhaar privacy issues (feasible) with handling strangers or malicious attackers. Our approach will show that it does not reveal any information of any user to anyone, i.e. it provides (complete) high uncertainty and high anonymity in a system. Also in extension of this work, more robustness tests like including of cryptographic applications will be done in future. By the way, collection, and protection of captured biometric information (stored with UIDAI) is another area to address. In summary, we need fast accessing (with low cost, and less time) with technology improvements/advances in biometrics usage. But in real scenario, things are different, i.e. installation of biometric systems put a financial burden to organisation (via new equipment). So, we hope Real Virtual ID can provide such services without revealing information of users and avoiding irrelevant cost to implement, it will attract many users to use this scheme (also to prevent ID misuse by ID theft). On the other side, several devices are being developed and sold in the market (openly) for public use. Finally, the Real Virtual ID is a demonstration of

ID management and service. This idea can be extended and exploited to any other system, which requires anonymity and accountability.

Hence, this section discusses an open discussion for Aadhaar and put several future perspectives regarding Aadhaar. In summary, we reached to a conclusion that giving our fingerprint or sharing Aadhaar number for getting benefits from several welfare schemes does not reveal any kind of your personal information to other party/organisations. As discussed above, we (all) are sharing our iris scan as face lock and fingerprints to our smart devices, whereas this information is collected by the respective mobile company at their server. Remember always, the user is always a responsible person for leaking his/her identity or personal information to malicious users. Privacy preservation to users can be provided by proving higher unlinkability or anonymity to user's information or building trust among people. Now, the next section will conclude this work in brief.

5 Conclusion

We need Real Virtual ID to provide anonymity (also unlinkability to user's information) to different organisations to users. This RVID cannot be generated by the user without the presence of Internet. This generated RVID cannot be used or tampered by agencies like AUA/KUA/ASA for de-duplication. Data de-duplication reduces storage costs and processing overhead, i.e. accessing cost and maintaining cost. In this approach, redundant data blocks are removed and replaced with pointers to the unique data copy for providing fast accessing to particular information. In the past research/literature work, we found that the Aadhaar number, which is a single national identifier that is supposed to work across application domains, makes individuals vulnerable to privacy breaches. For that government initialised a concept of using VID in the place of using their Aadhaar number. But using VID was an easy task but every time, it was a burden to Aadhaar database/system. So this work enhances and removes this issue by providing an extra layer of security, i.e. 16-digit unique token to every Aadhaar card holder. In future as the extension of our proposed work, we will come back with experimental results, i.e. with respect to this approach. Hence, in simple terms, a good design alteration can make Aadhaar system safe.

Privacy still remains a point of paradox and in the absence of concrete privacy laws, citizens might be subjected to mass surveillance in the name of national security. In current Aadhaar database, the biggest threat to user's privacy comes from insider attackers/insider leaks. The current Aadhaar technology architecture has tried to overcome such attacks but still does not have designed a strong system, which can protect such insider leaks. Hence, we/future researchers and the government still require so much research to be done with respect to protect Aadhaar database. Also, the government needs much more dedicated, informed and comprehensive security

policies and accelerated efforts to realise Aadhaar's full effectiveness. Thus, with appropriate measures on the security front, Aadhaar can be associated with numerous benefits like a cashless society, reduction of voter fraud and legitimate allocation of subsidies.

Author Contributions Amit Kumar Tyagi conceived of the work, and drafted the manuscript, whereas G. Rekha designed the schemes. Amit Kumar Tyagi and Sreenath Niladhuri contributed to the original ideas and scheme design.

References

1. <https://simplybanking.wordpress.com/2013/07/14/the-actual-uidai-aadhaar-number-is-11-digits-long-and-not-12-digits/>
2. <https://economictimes.indiatimes.com/wealth/personal-finance-news/aadhaar-everything-you-need-to-know-about-it/articleshow/60173210.cms>
3. <https://scroll.in/article/864698/explainer-what-is-virtual-id-and-how-is-it-different-from-aadhaar>
4. <https://www.paisabazaar.com/aadhar-card/aadhaar-authentication/>
5. <https://uidai.gov.in/authentication/authentication-partners/user-agency.html>
6. <https://authportal.uidai.gov.in/home-articles?urlTitle=requesting-entities&pageType=partners>
7. Ríha, Z., Matyáš, V.: Biometric Authentication Systems, FI MU Report Series (2000)
8. Tyagi, A.K., Sreenath, N., Priya, R.: Never trust anyone: trust-privacy trade-offs in vehicular ad hoc network. *Br. J. Math. Comput. Sci.* **19**(6), 1–23 (2016, November). ISSN: 2231-0851
9. Tyagi, A.K., Rekha, G., Sreenath, N.: Is your privacy safe with Aadhaar?: an open discussion. In: Proceedings of Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC) 2018, India (2018)

A Voting-Based Sentiment Classification Model



Dhara Mungra, Anjali Agrawal and Ankit Thakkar

Abstract Sentiment analysis is used to depict sentiments present in the text structures, including news, reviews, and articles, and classify them as positive, or negative. It has gained significant attention due to the increase in individuals utilizing social media platforms to express sentiments about organizations, products, and administrations. Many methods are being devised to improve the efficacy of automated sentiment classification. The study proposes a voting-based ensemble model Majority Voting (MV) using five supervised machine learning classifiers named Logistic Regression (LR), Support Vector Machine (SVM), Artificial Neural Network (ANN), Decision Tree (DT), and Random Forest (RF) as base classifiers and a majority voting rule-based mechanism to get the final prediction. The performance of the proposed method is assessed using minimum, maximum, mean, and median values of precision, recall, f -score, and accuracy. The results of 900 values of the classification accuracy (3 datasets * 6 (classification methods) * 10 data subsets (k -fold cross-validation for $k = 10$) * 5 runs), indicates that the proposed approach outperforms the individual classifiers in majority of the cases.

Keywords Sentiment analysis · Preprocessing · Binary class · Majority voting · Different domains

1 Introduction

Developments in social media and e-commerce platforms have resulted in a growth in the number of people expressing their opinions about a movie, product, food

D. Mungra (✉) · A. Agrawal · A. Thakkar
Institute of Technology, Nirma University, Ahmedabad 382481, Gujarat, India
e-mail: 15bce031@nirmauni.ac.in

A. Agrawal
e-mail: 15bce008@nirmauni.ac.in

A. Thakkar
e-mail: ankit.thakkar@nirmauni.ac.in

item, airline, restaurant, event, or any subject of general interest on the internet [7]. Sentiment analysis aims at analyzing these opinions, containing a lot of useful information relating to consumers, products or any new policy, and classify them into positive or negative sentiment class [8]. This analysis can be performed manually or automatically. Automatic sentiment analysis and classification of the data is gaining popularity due to the increasing number of people expressing their opinions on Internet [7]. Machine learning classifiers named Naïve Bayes (NB) [7], Logistic Regression (LR) [4], Support Vector Machine (SVM) [4, 12], Artificial Neural Network (ANN) [7, 8], Decision Tree (DT) [7], and Random Forest (RF) [2], are widely used by researchers for the task of sentiment analysis due to their efficacy [7]. Recently, ensemble models such as Bayesian Ensemble Learning, Majority voting (MV), Bagging, etc., that takes advantage of multiple base classifiers to predict class label of test dataset more accurately are gaining a lot of popularity [9].

This study proposes the use of ensemble model MV using five supervised machine learning classifiers LR, SVM, ANN, DT, and RF as base classifiers for the sentiment analysis of text using three balanced binary datasets from different domains having varying features. For the experimentation, the text is preprocessed using contraction replacement, text cleaning, lowercasing, tokenization, stop words removal, and word stemming, followed by feature extraction and selection using Bag of Words (BOW) model and Information Gain (IG), respectively. The experiments are carried out using k -fold cross-validation for $k = 10$ for five iterations [3]. The performance of MV is compared against the performance of base classifiers named LR, SVM, ANN, DT, and RF using precision, recall, f -score, and accuracy as various performance measures.

2 Related Work

In the literature, a lot of work has been presented to address the topic of sentiment analysis. The study [1] presents an empirical comparison of the performance of 16 preprocessing techniques such as stopwords removal, word stemming, contractions replacement, etc., for the sentiment analysis of text data. Results from the experiments showed that the methods such as lemmatization, eliminating numbers and replacing contractions improves the performance of the classifiers. The paper [2] compares the performance of four popular machine learning classifiers named RF, Multinomial NB, SVM, and Stochastic Gradient Descent for cross-domain and single-domain sentiment analysis. Experimental results indicated that SVM-RBF showed the slowest performance among the four classifiers used while Multinomial NB and Stochastic Gradient Descent showed consistent performance. In the study [7], the performance of four popular feature selection algorithms named IG, gain ratio, document frequency, and CHI statistics, and five popular machine learning classifiers named DT, NB, radial basis function neural network, SVM, and k -nearest neighbor for multi-class sentiment classification are compared and evaluated. Simulation results showed that SVM outperforms the other four classifiers used. The paper [8]

analyses, assesses, and compares the performance of SVM and ANN classifiers for the task of sentiment classification. ANN classifier performed comparatively better than SVM with less training time. The study proposes [9], the use of three ensemble models named MV, weighted voting, and hybrid approach using NBSVM—a variant of SVM model, CNN, and lexicon-based classifiers as a base classifiers for Cantonese sentiment analysis. Results indicated that weighted voting outperformed other two ensemble models and three classifiers used individually.

This study proposes a voting-based ensemble model MV using five machine learning classifiers named LR, SVM, ANN, DT, and RF as base classifiers. Performance of LR, SVM, ANN, DT, RF, and MV is evaluated and compared using minimum, maximum, mean, and median, of precision, recall, *f*-score, and accuracy obtained after performing five runs of each algorithm with different random seed values for three balanced binary datasets from different domains having varying features. The rest of the paper is organized as follows: The proposed approach is discussed in Sect. 3, Experimental setup and performance analysis are presented in Sect. 4, and Concluding remarks are given in Sect. 5.

3 Proposed Approach

Sentiment analysis is the task of recognizing and classifying a given text into positive or negative sentiment class computationally [7]. Preprocessing of the text is the first step while performing sentiment analysis of text data. Text preprocessing begins with contraction replacement to replace strings of the forms “haven’t” with “would not” [11] followed by text cleaning to remove punctuation marks, numbers, extra characters, and emoticons [1]. The text is converted to lowercase to make it more uniform, and tokens are then generated from this text [11]. Tokenization is followed by stopwords removal in which all the stopwords except “not”, “no”, and “never” are removed as these words help the classifier in identifying the negative sentiments present in the text [11], and word stemming for combining the words having the same base [11]. Preprocessed text is transformed into vectors after extracting meaningful features from this text using simple yet effective BOW model [13]. Relevant and important features are selected from these extracted features using the most widely used feature selection algorithm IG [7].

Selected features are fed to classifiers for training which classify a new observation into positive or negative sentiment class [5]. For the experimentation, we have used five popular and effective machine learning classifiers named LR, SVM, ANN, DT, and RF. LR is a popular linear, probabilistic classifier that performs classification by fitting the input data into a logistic function [4]. The probability that a particular input sample belongs to a particular class is calculated using the distance of input from the hyperplane [4]. SVM is one of the most widely used supervised machine learning technique that separate two classes using optimal hyperplane as a decision boundary[4]. These boundaries are defined using a kernel function [4]. For binary classification, a linear kernel is used [4]. ANN is the structural and functional imi-

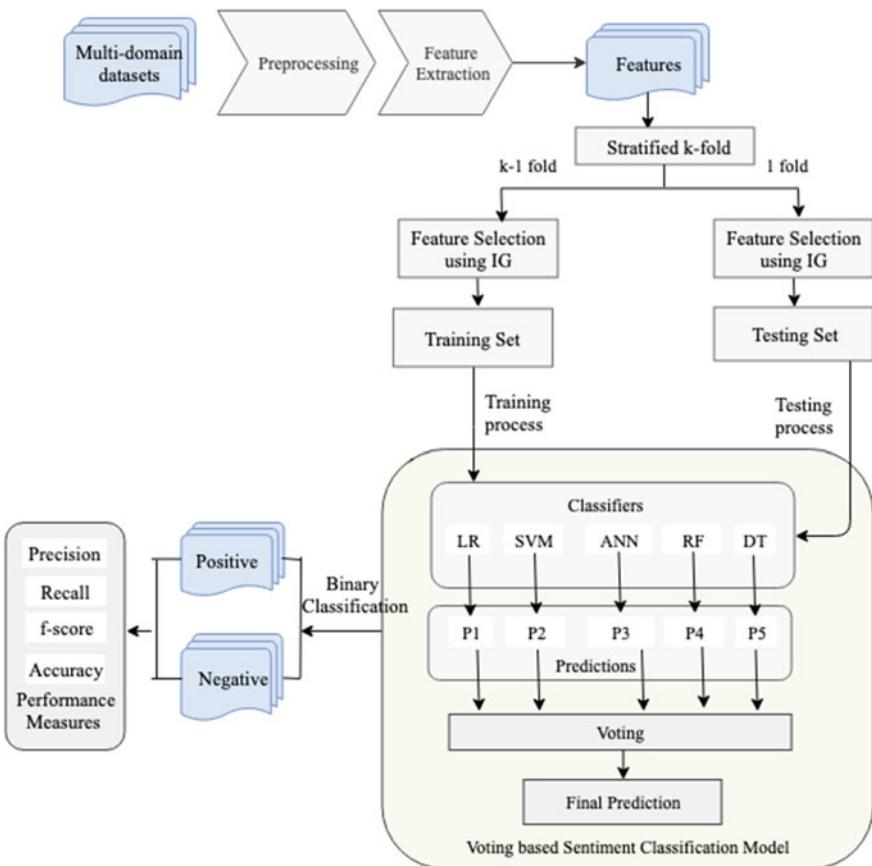


Fig. 1 Voting-based sentiment classification using binary datasets for various domains

tation of human brain [10]. It is one of the simplest and effective machine learning classifiers comprising of neurons and three types of layers: an input layer, number of hidden layers, and an output layer [8]. DT is one of the most common algorithms that predicts the class of data samples using a tree constructed from training data [7]. Multiple DTs are combined into single RF to produce more accurate results [2]. Decorrelated decision trees generated from random data in the dataset are used to produce the aggregated result [2].

This study proposes the use of popular ensemble model MV using five machine learning classifiers named LR, SVM, ANN, DT, and RF as base classifiers for the sentiment analysis of text data. In MV, these classifiers are combined to predict the class labels for the test samples [9]. Class value with more than half of the total votes from the individual classifiers is selected [9]. Experiments are carried out using k -fold cross-validation for $k = 10$. Each of these experiments is performed for five iterations, and the process is outlined in Fig. 1.

4 Experimental Setup and Performance Analysis

We have performed experimentation on macOS High Sierra 64-bit operating system running on a machine with RAM size of 4 GB and Intel(R) Core(TM) i5 processor with the clock frequency of 1.6 GHz. For the experimentation, we have used three balanced binary datasets [6] of same varying number of features. Results from the experiments are obtained in the form of confusion matrix shown in Table 2. This confusion matrix is used to classify the comment or review from the binary dataset into one of the four categories: True Positive (TP), True Negative (TN), False Positive (FP), or False Negative (FN) [4]. Performance of five individual machine learning classifiers named LR, SVM, ANN, DT, and RF, and the ensemble model MV is evaluated using minimum, maximum, mean, and median values of precision, recall, *f*-score, and accuracy. Values of precision, recall, *f*-score, and accuracy are calculated using Eqs. 1–4 respectively.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$f\text{-score} = \frac{2 * precision * recall}{precision + recall} \quad (3)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

Table 1 gives the minimum, maximum, mean, and median of the values of precision, recall, *f*-score, and accuracy obtained after five runs with different random seed values. The results highlighted in boldface indicate the best results in Table 1. Ensemble model MV outperforms LR, SVM, ANN, DT, and RF classifiers used individually in terms of accuracy in almost all the cases as can be observed from Table 1. Using majority voting rule, it can be concluded that proposed approach MV outperforms to the rest of five classifiers for recall, accuracy, and *f*-score as far as mean and median values are used to measure the performance of the classifiers. However, SVM outperforms to MV for the mean value of the precision while MV outperforms for the median value of the precision. This improvement is observed as each classifier gives its opinion about each of the data sample and MV concludes the polarity of the data sample based on the majority voting received for a specific data sample (Table 2).

Table 1 Performance analysis of MW for different test datasets using k-fold cross-validation for $k=10$

Approach	Precision			Recall			f -score			Accuracy		
	Min	Max	Mean	Median	Min	Max	Mean	Median	Max	Mean	Median	Mean
<i>IMDb dataset [6]</i>												
LR	79.02	80.57	79.65	79.54	77.20	78.80	77.92	77.60	77.99	79.54	78.60	78.35
SVM	78.90	81.46	80.06	79.68	75.80	78.40	77.20	77.00	77.16	79.60	78.49	78.24
ANN	76.22	77.98	77.03	77.08	78.80	80.60	79.44	79.20	77.51	78.64	78.03	77.90
RF	73.74	75.44	74.84	75.08	72.40	77.40	75.04	74.80	73.20	76.05	74.58	74.36
DT	70.45	72.66	71.02	70.72	72.80	75.20	73.92	73.80	71.82	72.42	72.11	72.05
Voting	79.14	80.30	79.70	79.83	78.60	81.60	79.76	79.40	79.13	80.72	79.54	79.35
<i>Amazon dataset [6]</i>												
LR	82.63	83.86	83.13	83.10	79.20	80.00	79.56	79.60	80.65	81.33	81.15	81.25
SVM	83.56	85.10	84.40	84.52	79.60	81.60	80.32	80.00	81.77	82.49	82.13	82.04
ANN	79.97	82.46	81.38	81.76	78.80	81.80	80.88	81.20	80.08	81.84	80.88	80.38
RF	80.76	83.71	82.67	82.97	76.20	79.40	78.20	79.00	79.44	80.71	79.92	79.62
DT	77.57	80.20	78.71	78.39	80.20	83.20	81.12	80.60	79.06	80.02	79.57	79.72
Voting	83.34	84.45	83.81	83.85	80.00	82.80	81.36	81.60	81.46	83.39	82.30	82.20
<i>Yelp dataset [6]</i>												
LR	81.87	82.68	82.19	82.12	79.00	80.80	79.72	79.60	80.33	81.08	80.71	80.72
SVM	81.56	83.74	82.51	82.42	79.40	80.80	80.20	80.40	80.48	81.81	81.15	81.39
ANN	80.55	81.68	81.27	81.30	80.20	82.00	81.04	80.28	81.51	81.03	81.10	80.30
RF	80.82	84.26	82.31	81.60	75.60	78.80	76.84	77.00	77.97	80.00	79.02	79.00
DT	77.00	78.93	78.05	78.36	79.20	81.80	80.24	79.80	78.56	79.14	78.81	78.82
Voting	81.93	84.78	83.34	83.34	79.40	81.60	80.20	80.00	80.45	82.37	81.50	80.80

Table 2 Confusion matrix

	Predicted negative comment	Predicted positive comment
Actual negative comment	True negative (TN)	False positive (FP)
Actual positive comment	False negative (FN)	True positive (TP)

5 Concluding Remark

The study proposes a voting-based ensemble model MV using five supervised machine learning classifiers named LR, SVM, ANN, DT, and RF for the balanced binary datasets of different domains and features. Performance of MV is compared against the performance of individual classifiers LR, SVM, ANN, DT, and RF, to prove the efficacy of the MV model. Results from the experiments performed using 10fold cross-validation, showed that MV gives higher or comparable mean and median values of all the performance measures named precision, recall, f -score, and accuracy.

References

1. Angiani, G., Ferrari, L., Fontanini, T., Fornacciari, P., Iotti, E., Magliani, F., Manicardi, S.: A comparison between preprocessing techniques for sentiment analysis in twitter. In: KDWeb (2016)
2. Aziz, A.A., Starkey, A., Bannerman, M.C.: Evaluating cross domain sentiment analysis using supervised machine learning techniques. In: Intelligent Systems Conference (IntelliSys), 2017. pp. 689–696. IEEE (2017)
3. Borra, S., Ciaccio, A.D.: Measuring the prediction error. A comparison of cross-validation, bootstrap and covariance penalty methods. Comput. Stat. Data Anal. **54**(12), 2976–2989 (2010)
4. Castro, D.W., Souza, E., Vitório, D., Santos, D., Oliveira, A.L.: Smoothed n-gram based models for tweet language identification: a case study of the brazilian and european portuguese national varieties. Appl. Soft Comput. **61**, 1160–1172 (2017)
5. Jha, V., Savitha, R., Shenoy, P.D., Venugopal, K., Sangaiah, A.K.: A novel sentiment aware dictionary for multi-domain sentiment classification. Comput. Electr. Eng. (2017)
6. Kotzias, D., Denil, M., De Freitas, N., Smyth, P.: From group to individual labels using deep features. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 597–606. ACM (2015)
7. Liu, Y., Bi, J.W., Fan, Z.P.: Multi-class sentiment classification: the experimental comparisons of feature selection and machine learning algorithms. Expert. Syst. Appl. **80**, 323–339 (2017)
8. Moraes, R., Valiati, J.F., Neto, W.P.G.: Document-level sentiment classification: an empirical comparison between svm and ann. Expert. Syst. Appl. **40**(2), 621–633 (2013)
9. Ngai, E., Lee, M., Choi, Y., Chai, P.: Multiple-domain sentiment classification for cantonese using a combined approach (2018)
10. Paliwal, P., Kumar, D.: Abc based neural network approach for churn prediction in telecommunication sector. In: International Conference on Information and Communication Technology for Intelligent Systems, pp. 343–349. Springer (2017)
11. Symeonidis, S., Effrosynidis, D., Arampatzis, A.: A comparative evaluation of pre-processing techniques and their interactions for twitter sentiment analysis. Expert. Syst. Appl. (2018)

12. Tellez, E.S., Miranda-Jiménez, S., Graff, M., Moctezuma, D., Siordia, O.S., Villaseñor, E.A.: A case study of spanish text transformations for twitter sentiment analysis. *Expert. Syst. Appl.* **81**, 457–471 (2017)
13. Williams, L., Bannister, C., Arribas-Ayllon, M., Preece, A., Spasić, I.: The role of idioms in sentiment analysis. *Expert. Syst. Appl.* **42**(21), 7375–7385 (2015)

Dynamic Performance of Grid-Connected Wind Farms with and Without UPFC: A Case Study on Ashagoda Wind Farm



Asefa Sisay and Vibhu Jately

Abstract Voltage variations are a major concern in grid-connected wind turbines systems. In Ashagoda wind farm, different voltage quality problems such as voltage sag, voltage dip, and voltage fluctuations have been observed for causes like loss of system voltage, short circuit faults, and tripping of wind turbines that resulting in a huge power loss. Under three-phase short circuit fault, the voltage at the point of common coupling (PCC) instantly drops below 80% of its nominal value which results in instability in the rotor speed. In this paper, a unified power flow controller (UPFC) is interconnected at the PCC to enhance the dynamic response of ac grids and also improves the transmission capability of the system. The system is modeled and simulated in Dig Silent Power Factory. The results of simulation clearly illustrate that UPFC can improve voltage regulation at the point of common coupling and maintain a reliable connection between the wind turbine and the grid under minor levels of voltage fluctuation on the grid side.

Keywords WTG · DFIG · FACT device · UPFC

1 Introduction

Technology for harnessing energy from the Sun and from the wind have been developed over the last three decades in a significant way. Energy from the Sun is available in abundance in tropical regions, but it is available only during the day. On the other hand, in general, wind in many locations is stronger during the night in comparison to day time. There have been incentives by international agencies as well as by the national governments for setting up generating plants base on wind turbines. The size of the turbines have been increasing and locations having potential for setting these

A. Sisay · V. Jately (✉)
Department of ECE, KIoT, Wollo University, Kombolcha, Ethiopia
e-mail: vibhujately@gmail.com

A. Sisay
e-mail: asefasisay581@gmail.com

turbines up are being identified. The ratio of wind-based generating plants to the total installed capacity is rising, and in some countries in Europe, it is now above 15% [1]. Increase in this ratio presents a number of challenges to the engineers managing the power system. These are related to system security, power imbalance, management of reserve capacity, system voltage control, power quality, i.e., degree of harmonics in supply, flicker due to frequency changes, voltage dip, fault ride through capability, and stability of the power system [2–4].

It is worth noting that most of the generation in conventional power plants such as hydroelectric, nuclear, or thermal is through synchronous generators. However, almost all wind turbines are coupled to induction generators. In grid-connected plants, synchronous generators are able to assist in maintaining the node voltages in power system within the prescribed limits as these generators are able to assure reactive power balance by adjusting the corresponding excitation levels [5–7]. This reactive power balancing capability is missing in case of induction generators. Therefore, the problems of voltage dips or sag, which may occur due to changes in load demand under normal and steady conditions, or on occurrence of faults in the system under dynamic conditions, have to be addressed in grid-connected wind plants by external means. Flexible AC Transmission Systems (FACTS) devices such as the Static Synchronous Compensator (STATCOM), Static Series Synchronous Compensator (SSSC), and the Unified Power Flow Controller (UPFC) that are often being used in power systems to provide flexible power flow control can help in mitigating these problems associated with on-grid wind plants [8–10].

In this paper, the effect of connecting a UPFC on an existing grid-connected wind power plant is analyzed. The designed UPFC, when connected in the system considerably improves the dynamic response of Ashagoda wind power system. The mathematical modeling of the entire system is done in MATLAB/Simulink environment. The system is simulated under various test cases. The results obtained clearly demonstrate the effectiveness of the proposed addition.

2 Ashagoda Wind Power System

Ashagoda wind farm consists of five wind turbines: two of them have a maximum generating capacity of 1 MW (WT_1 and WT_2) the remaining three have a maximum generating capacity of 1.67 MW (WT_3, WT_4, and WT_5). All the wind turbines are connected to a 33 kV distribution system, which then exports power to a 230 kV grid through an 18 km, 33 kV feeder as shown in Fig. 1.

The Ashagoda wind power system suffers from large voltage variations and exhibits poor dynamic response under fault conditions. To overcome these drawbacks, the authors have designed a UPFC that interconnects the 33 kV distribution system with the 230 kV grid as shown in Fig. 2.

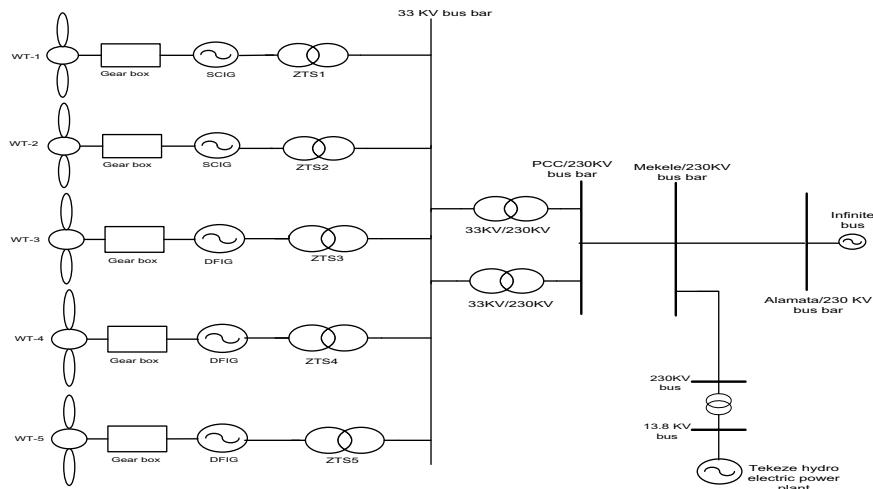


Fig. 1 Ashagoda wind power system

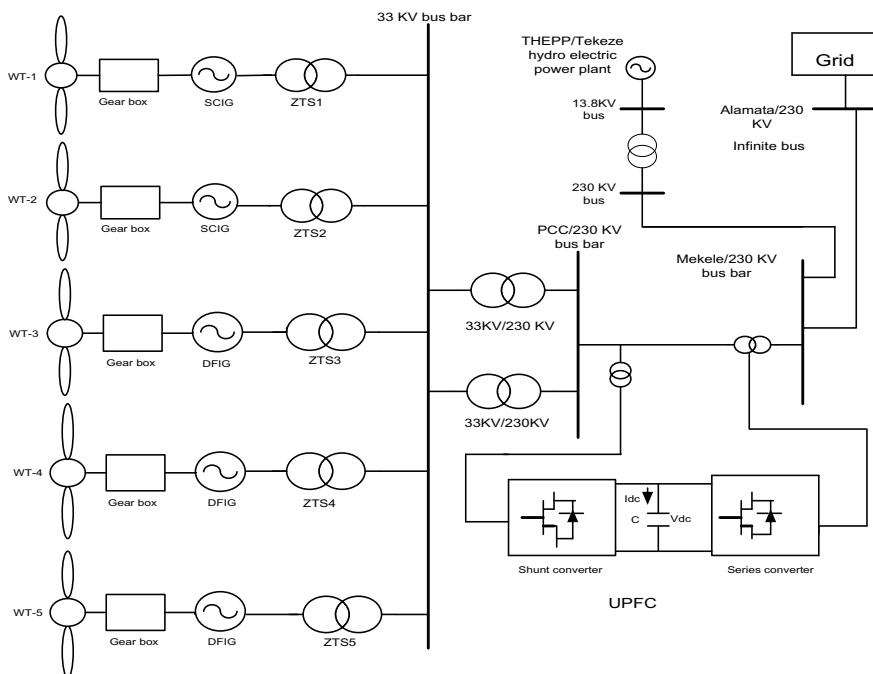


Fig. 2 Ashagoda wind power system with UPFC

3 Simulation Results

The stability analysis of Ashagoda wind farm is done for two test cases, with and without the UPFC. Both cases are simulated in Dig Silent Power Factory environment. In the simulation, the grid sides are represented by external networks. Grid side (external grid) represents the Ethiopian Electric Power Corporation (EEPCO) grid at Alamata. The external networks are provided with parameters of short circuit current and R/X ratio taken from the actual EEPCO grid to create an operational resemblance to the grid. The analyses are carried out under normal steady-state conditions as well as under fault conditions.

The designed UPFC enhances the dynamic responses of both ac grids, i.e., wind turbine generator and the external grid by:

- Improving the terminal voltage of the wind turbine generator.
- Improving the voltage magnitude at the Point of Common Coupling.

Case 1: Without UPFC

The first case analyzes the dynamic response of the system without UPFC.

(i) Steady-state operating conditions:

- a. The collector point bus voltages under steady-state operating condition are shown in Fig. 3.

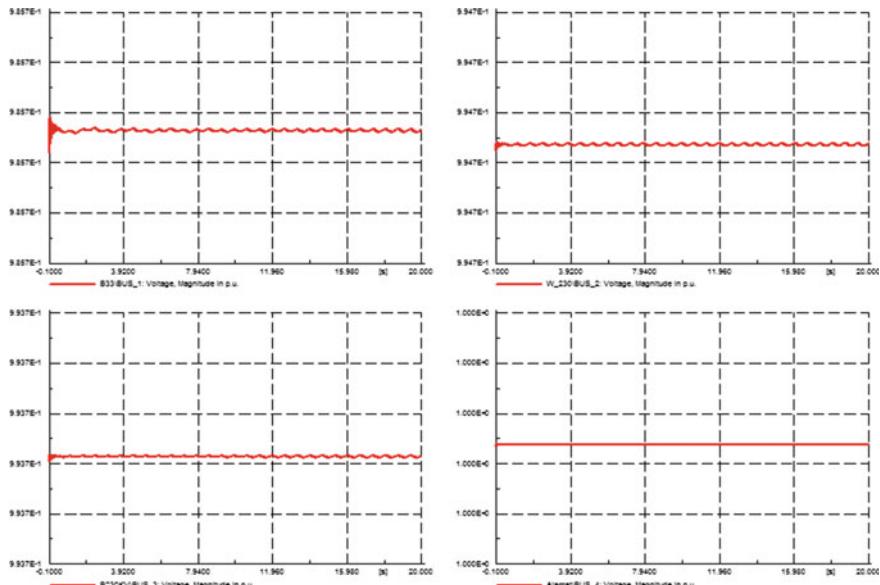


Fig. 3 Collector bus/33 kV bus, PCC/230 kV bus, Mekele/230 kV bus and Alamata/230 kV bus voltage magnitude in per unit under steady-state operating conditions without UPFC

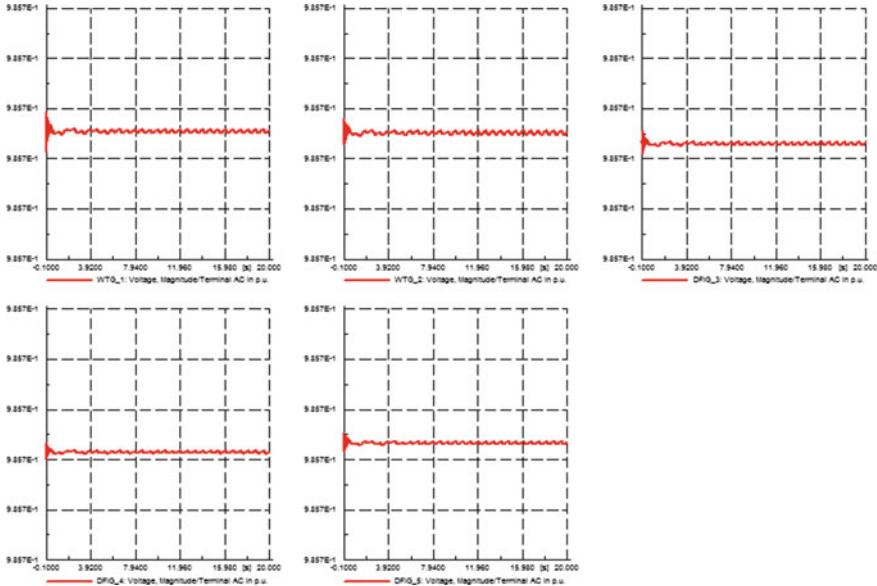


Fig. 4 Voltage magnitude in per unit of WTG under steady-state operating condition

- b. The terminal voltages of the wind turbines are shown in Fig. 4.
- (ii) **Abnormal operating conditions (three-phase short circuit fault):** The effect of a three-phase high impedance short circuit fault at the collector bus (33 kV) is studied. The ground fault is initiated at $t = 0$ s and cleared at $t = 0.2$ s.
 - a. The collector bus voltages under fault condition are shown in Fig. 5.
 - b. The terminal voltages of the wind turbine generator under fault conditions are shown in Fig. 6.

Case 2: With UPFC

The improved dynamic performance of UPFC-based grid-connected Ashagoda wind farm is observed under two conditions.

- (i) **Steady state operating conditions:**
 - a. The collector bus voltages at PCC under steady-state operating conditions are shown in Fig. 7.
 - b. The per unit terminal voltages of wind turbines with UPFC is shown in Fig. 8.
- (ii) **Abnormal operating conditions (three-phase short circuit fault):** The designed UPFC improves the dynamic response of the system under fault conditions. Under fault conditions, the per unit response of both the voltages: wind turbine terminal voltage and the external grid voltage are improved.

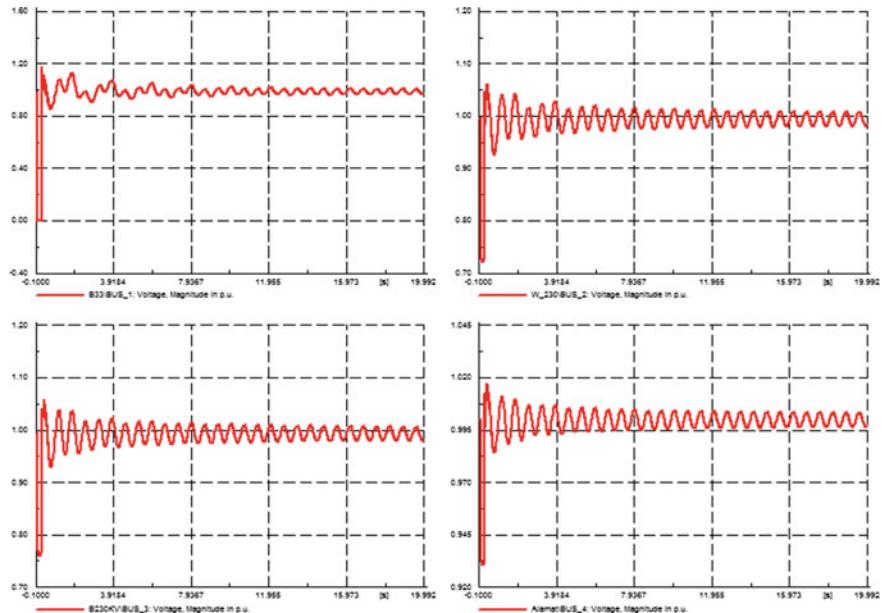


Fig. 5 Collector bus/33 kV bus, PCC/230 kV bus, Mekelle/230 kV bus, and Alamata/230 kV bus voltage magnitude in per unit under three-phase short circuit fault without UPFC

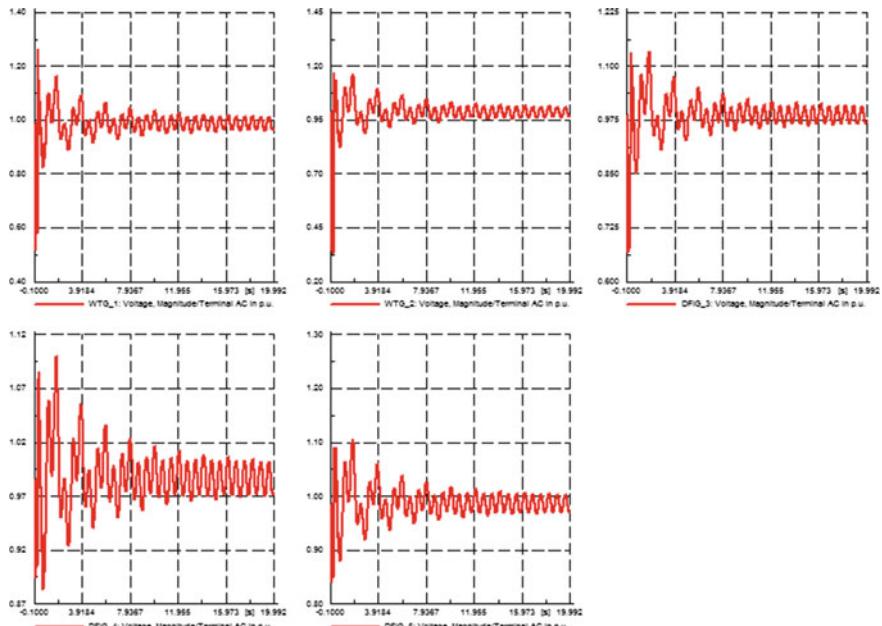


Fig. 6 WTG terminal voltages in per unit during three-phase short circuit fault without UPFC

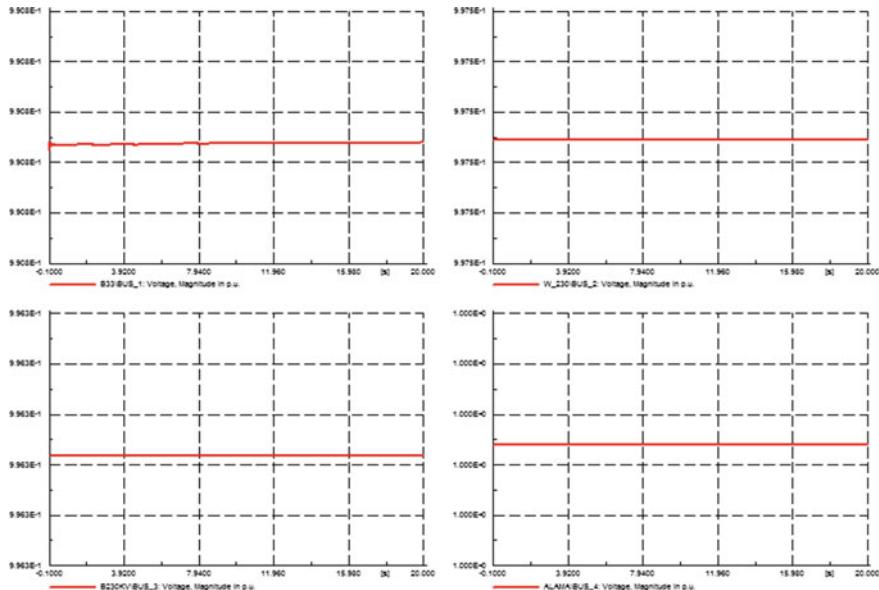


Fig. 7 Collector bus/33 kV bus, PCC/230 kV bus, Mekele/230 kV bus, and Alamata/230 kV bus voltage magnitudes in per unit under steady-state operating conditions with UPFC

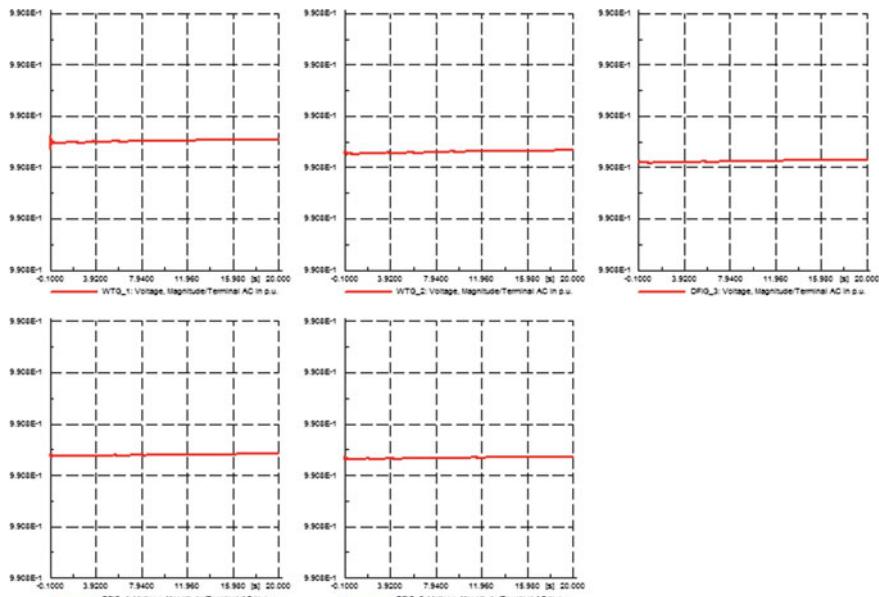


Fig. 8 WTG terminal voltage magnitudes in per unit under steady-state operating conditions with UPFC

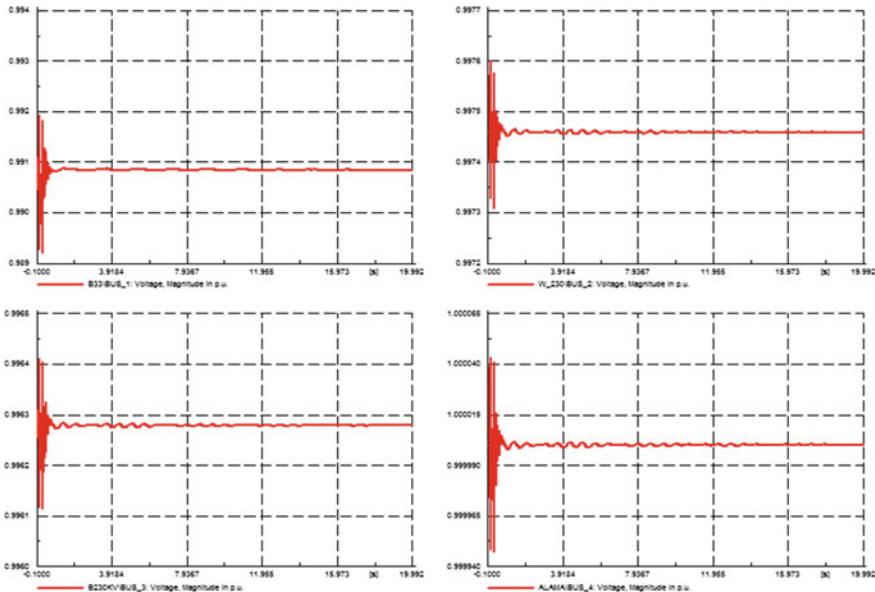


Fig. 9 Collector bus/33 kV bus, PCC/230 kV bus, Mekele/230 kV bus, and Alamata/230 kV bus voltage magnitude in per unit under three-phase short circuit fault with UPFC

- The collector bus voltages at PCC under steady-state operating conditions are shown in Fig. 9.
- The per unit terminal voltages of wind turbines with UPFC are shown in Fig. 10.

4 Conclusion

Results of the simulation studies conducted here have shown that the additional voltage/var support provided by an external device, called UPFC, can significantly improve the wind turbines' fault recovery by quickly restoring voltage characteristics to within the permissible limits. The interconnection of wind farms to weak grids, in general, poses a concern about the safety of wind turbine generators. Wind turbines connected to weak grids may lead to frequent faults, abnormal grid conditions, variation in frequency magnitude of grid voltage leading to tripping of relays and in acute cases loss of generation. Various simulations carried out in the previous section, establish that the dynamic performance of wind farms in a power grid is improved by the application of a UPFC. This work shows that the Ashagoda wind farm dynamic performance can be enhanced by improving voltage profiles both at point of common

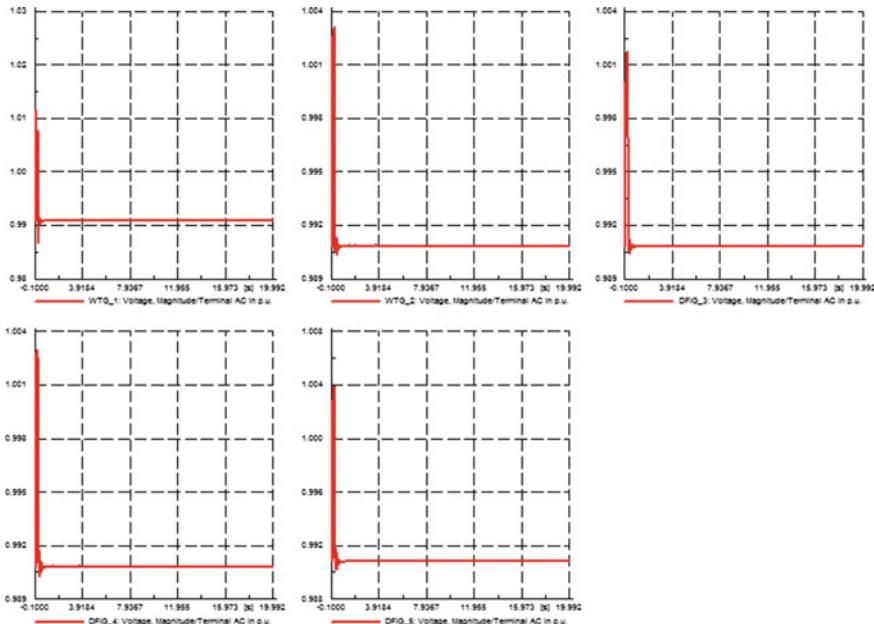


Fig. 10 WTG terminal voltages in per unit during three-phase short circuit fault with UPFC

coupling and at wind turbine generator terminals by adding a UPFC on the existing system. The performance analysis of all the clusters at Ashagoda wind farm will be done in the future.

References

1. Tapia, A., Tapia, G., Ostolaza, J.X., Saenz, J.R.: Modeling and control of a wind turbine driven doubly fed induction generator. *IEEE Trans. Energy Convers.* **18**(2), 194–204 (2003)
2. Mohammadi, M., Gitizadeh, M., Roosta, A.: Dynamic stability improvement of a power system incorporating DFIG wind power plant using optimized control parameters of a SVC. In: Power Engineering and Optimization Conference (PEDCO) Melaka, Malaysia, 2012 IEEE International 2012 June 6, pp. 416–421. IEEE (2012)
3. Sun, T.: Power Quality of Grid-Connected Wind Turbines with DFIG and Their Interaction with the Grid. Institute of Energy Technology, Aalborg University (2004, May)
4. Molinas, M., Vazquez, S., Takaku, T., Carrasco, J.M., Shimada, R., Undeland, T.: Improvement of transient stability margin in power systems with integrated wind generation using a STATCOM: an experimental verification. In: 2005 International Conference on Future Power Systems, 2005 November 18, 6-pp. IEEE (2005)
5. Yagami, M., Hasegawa, T., Tamura, J.: Transient stability assessment of synchronous generator in power system with high-penetration photovoltaics. In: 2012 15th International Conference on Electrical Machines and Systems (ICEMS), 2012 October 21, pp. 1–6. IEEE (2012)
6. Gomez-Exposito, A., Conejo, A.J., Canizares, C.: Electric Energy Systems: Analysis and Operation. CRC Press (2018)

7. Tan, K., Islam, S.: Optimum control strategies in energy conversion of PMSG wind turbine system without mechanical sensors. *IEEE Trans. Energy Convers.* **19**(2), 392–399 (2004)
8. Firouzi, M., Gharehpetian, G.B.: LVRT performance enhancement of DFIG-based wind farms by capacitive bridge-type fault current limiter. *IEEE Trans. Sustain. Energy* **9**(3), 1118–1125 (2018)
9. Krishnan, V.V., Srivastava, S.C., Chakrabarti, S.: A robust decentralized wide area damping controller for wind generators and FACTS controllers considering load model uncertainties. *IEEE Trans. Smart Grid* **9**(1), 360–372 (2018)
10. Öztürk, A., Döşoğlu, K.: Investigation of the control voltage and reactive power in wind farm load bus by STATCOM and SVC. In: International Conference on Electrical and Electronics Engineering, 2009. ELECO 2009, 2009 November 5. pp. I–60. IEEE (2009)

A Study on Agent-Based Web Searching and Information Retrieval



Urvi Mitra and Garima Srivastava

Abstract In this day and age of information overload, it becomes difficult to filter information that is relevant to our needs. World Wide Web is the largest database of information available to us. It stores information using the hypertext paradigm, i.e., interlinking web pages through hyperlinks, which users can click on to access related information. An agent acting on behalf of humans, can make the task of sifting through information to find what we need easier for us. This paper focuses on the application of intelligent agents in the field of web browsing and searching, mostly web spidering, indexing, and retrieval of information most relevant to us, based on keywords, from the vast database of knowledge available.

Keywords Agents · Information · Intelligent · Retrieval · Searching · Web

1 Introduction

1.1 *Information Retrieval*

Currently, information retrieval [1] is passive and it is based on searching static databases based on sorting documents [2] and web pages on the basis of factors like availability of search terms, their frequency, closeness to the beginning of the page, how close they occur in the document, etc. These criteria do not necessarily give the users the information they need. Therefore, a degree of personalization (customization according to a particular users' specific needs) is needed which renders better

U. Mitra (✉) · G. Srivastava
Amity University, Lucknow, India
e-mail: mitraurvi@gmail.com

G. Srivastava
e-mail: gsrivastava1@lko.amity.edu

and more useful search results. This is accomplished by the use of intelligent agents [3], specifically personal assistants. Three major information retrieval paradigms currently in use are statistical, semantic and contextual. Statistical information retrieval is done based on the statistical correlation of word count in documents and vector space models are made representing documents. Semantic information retrieval focuses on the underlying meaning of queries, uses Natural Language Processing and AI queries, and retrieves information based on term associations. Contextual information retrieval looks at the context and structure of the results returned by retrieval systems and encoded relationships among terms in documents, etc. Evaluation of the returned information is done by two standards—recall ratio, i.e., ratio of relevant documents returned to the number of relevant documents; and Precision ratio, i.e., the ratio of relevant documents returned to the number of documents retrieved.

1.2 *Introduction to Agent Concepts*

Agents are abstract concepts or concrete software programs that accomplish tasks that we require them to do, on behalf of us humans [4]. What separates agents from objects and robots is the element of autonomy they provide, which is why they are named “intelligent agents”. They make decisions and act out of their own discretion based on past experiences (for instance, case-based reasoning [5, 6]) and various evolutionary algorithms. Learning agents also learn from the feedback of their performance and evaluation criteria imposed on their results. They take actions to accomplish certain predefined goals and produce results based on criteria that define their appropriateness.

Components:

An agent has the following components:

- **Sensors** which detect and glean information about their environment
- **Environment** or the outside world or surroundings, in the context of the agent
- **Actuators** or agents parts that carry out the assigned tasks
- **Agent Program** that maps inputs to actions to be performed
- **Inputs or Percepts**, gleaned from the surroundings using sensors
- **Outputs**, based on predefined rules, i.e., programs and specified goal and criteria
- **Goal**, to be accomplished by the agent

Classification:

There are various categorizations of an agent [7]. A simple agent is based on simple if-else conditions preprogrammed into them. There is no deduction or reasoning involved, just a static database. Model-based agents are slightly better, in that they have a record of past cases encountered, which are used to find solutions to new problems. Learning agents learn and evolve through feedback and have an element of reasoning and Strong AI, helping them find solutions to unexpected problems.

These are the most complex of them all and are capable of making their own decisions leading to an effective and scalable accomplishment of the defined goal.

Agent Environments:

Environments can be deterministic, i.e., complete information could be available about them, based on which we can determine the effect of the agent's actions on it. On the other hand, there are non-deterministic environments, on whose state the effects of agents' actions cannot be determined. There are episodic environments, which can be divided into episodes, differentiated from each other, and non-episodic environments. In real life, as opposed to simulations, environments are non-episodic and non-deterministic.

Architectural Paradigm:

Agents are built based on three types of architectural paradigms. It could be an integrated part of an end program, with all the rules set. This requires the user to extensively set if-else conditions for the agent to act. The second is knowledge-based approach where the agent has domain-specific knowledge. In this approach, a lot of domain-specific knowledge is required. The third approach, namely learning approach removes the disadvantages of the previous two. In this approach, the agent learns from examples and cases and adds to its existing knowledge database.

2 A Web Searching Agent

Intelligent Agents can make use of the “spidering” technology, in use by most search engines out there [8]. A database of the users' likes and dislikes can be made using past browsing [9], history and information requested by the user. Based on this data, user profiles can be maintained [10]. Based on the keywords input by the user, information can be retrieved using Web Spiders that maintain most URLs on the web that are interrelated and have more or less the same type of information. Indexing [11] here is not a back of the book index but it is based on keywords and the type of information the web pages contain.

2.1 User Profile and Information Retrieval

User profiles contain information about users' interests and past searches. Based on that data, relevance may be assigned to the web pages output by search engines and web crawlers. A relevance function can be used to map the relevance of a web page to a user based on his/her user profile [12]. All web pages that cross a relevance threshold are output to the user. The user then gives implicit feedback by accessing and reading the web pages most valuable to him/her. This feedback provides further

information about the subjects the user may be interested in. Feedback could be in the form of explicit feedback and ratings by the user or implicit feedback like search results, i.e., web pages that are opened and accessed, time spent on them, printing the document or query result, saving its location or sampling passing it without doing anything. This feedback is processed and used to update the user profile, which is then again used for further searches. Such intelligent agents would be autonomous and capable of determining the relevance of a website on their own [13].

There are recommender systems—for instance, ones that give movie and book recommendations. There are also recommender systems [14] in the field of research that suggest useful and relevant material for the specified research topic. The search results containing the most relevant web pages are rendered by the intelligent agent and can be picked up by the user anytime he/she wants. This saves up time and energy for more important tasks.

2.2 *Retrieval Techniques*

Recommender systems use a technique called Automated Collaborative Filtering [15]. For instance, a movie recommendation system records user preference and maintains a user profile, based on information provided by the user, which is asked of them using relevant questions and web forms. The movie recommendations returned are rated by the user [16]. This feedback in the form of rating is used to further update the user profile and data about user preferences, which is used to recommend more movies. If a user likes a particular movie, more movies of a similar kind are searched from user profiles [17] and data of other users who have liked the same movie. This is based on the assumption that if a user likes a particular movie, he or she will like more movies of the same kind or genre. This, all users' data is an interrelated web of information that is crawled through, just like related web pages [18], to get useful information [19].

Another technique is content-based filtering where the content of different web pages is analyzed to group web pages with similar content together and a user who frequently browses a web page is recommended other web pages with similar content [20].

Thus, an autonomous agent is added to an existing search engine or retrieval algorithm [21] to improve retrieval performance based on measurable parameters. An agent, an information retrieval algorithm, a user preference algorithm and an existing search engine can be combined [22].

3 Some Web Searching Agents

3.1 *Web Crawlers*

Web crawlers or bots are also essentially intelligent agents that search and list viewed websites and URLs based on keywords input by the user. They list these websites and rate and rank them. Based on input search words, they retrieve those websites from something called crawler frontier—the list of URLs supposed to be visited by the bot. This contains all URLs and hyperlinks on all pages that are to be visited next. Web crawlers have certain policies to retrieve and display web pages. Web crawlers cannot retrieve the entire web, around 16% of the web. They have selection policies to select what to display based on page ranks, number of backlinks, breadth first search (which usually displays in order of page ranks because pages with higher page ranks have more links linking to them). They have revisit policies—since web pages keep updating themselves and crawlers need updates versions so these policies dictate how to visit them again after updating events—uniform policy visits all of them regardless of updating while proportional policy visits them in a frequency directly proportional to the rate of change. Web servers have implemented protocols to prevent access to certain parts by crawlers and some have the crawl-delay parameter specifying the minimum delay between the crawler accessing the same site.

3.2 *Internet Bots*

Internet Bots, web robots or simply, bots automate many of the repetitive tasks that are performed over the internet and perform them at a faster rate. They are used where the response rate required is faster than human response rate. An example of internet bot is a chatterbot. Chatterbots are used for services like reporting the weather, customer support in websites and in some cases are assigned characters that help the user navigate a web page easily, helping with instructions, etc.

3.3 *ChatterBot*

Chatterbots are conversational agents that talk to users in plain English. A user can ask them questions and chat with them, eliciting normal human-like, “intelligent” responses from them. Current chatterbots are based on weak AI, where pattern matching algorithms are used and they respond on the basis of static databases that store

keywords and relevant words, that are searched in user queries and responses formulated. Chatterbots are moving toward general conversations through learning in real time and evolutionary algorithms instead of just pattern matching, thus moving toward being truly “intelligent” and rational.

3.4 Examples

Some examples of web robots and crawlers that automatically traverse the web are

- Harvest, a part of the Harvest Project that runs from the University of Colorado, indexes topic-specific HTML objects
- CIGSearchbot is an example of collaborative information filtering, a multi-agent system based approach to information gathering
- Newt is an information filtering system which utilizes a society of agents that inhabit a user’s computer system. Each agent is a user profile that recommends documents that match itself and uses feedback from the user to improve their user profile
- AppleSearch is a commercial agent that searches information from AppleShare, Apple’s file sharing application
- NewWave is Hewlett-Packard’s agent feature and one of the oldest commercial agents. It customizes the user interface and automates simple tasks.
- Personal WebWatcher [23, 24] is a content-based web browsing agent that has two parts—a proxy server, that interacts with the user through the web browser and takes the user’s request and a learner that contains the user’s model, based on which the agent highlights hyperlinks relevant to the user and returns the web page with highlighted hyperlinks and notes
- PersonalSearcher [25] uses hierarchical clustering to fetch relevant and interesting search results for users based on user’s interests.

4 Review of Related Work

Web agents that use text-based machine relearning techniques [26] are listed in Tables 1, 2 and 3. Agents in Table 1 use content-based techniques, agents in Table 2 use collaborative techniques [27].

Table 1 Agents using content-based techniques

Agent	Organization	Goal	Publication
Antagonomy	NEC	Personalized newspaper	T. Kamba, H. Sakagami, and Y. Koseki, "Anatagonomy: A Personalized Newspaper on the World Wide Web," Int'l J. Human-Computer Studies, Vol. 46, No. 6, June 1997, pp. 789-803
Personal WebWatcher	CMU, IJS	Browsing WWW	D. Mladenic, Personal WebWatcher: Implementation and Design, Tech. Report WebWatcher IJS-DP-7472, Dept. of Computer Science, J. Stefan Inst., 1996; http://cs.cmu.edu/~TextLearning/pww
Syskill & Webert	UCI	Browsing WWW	M. Pazzani, J. Muramatsu, and D. Billsus, "Syskill & Webert: Identifying Interesting Web Sites," Proc. 13th Nat'l Conf. AI AAAI 96, AAAI Press, Menlo Park, Calif., 1996, pp. 54-61
WAWA	Wisconsin	Browsing WWW	J. Shavlik and T. Eliassi-Rad, "Building Intelligent Agents for Web-based Tasks: A Theory-Refinement Approach," Working Notes of Learning from Text and the Web, Conf. Automated Learning and Discovery (CONALD-98), Carnegie Mellon Univ., Pittsburgh, 1998; http://www.cs.cmu.edu/~conald/conald.shtml

(continued)

Table 1 (continued)

Agent	Organization	Goal	Publication
WebWatcher	CMU	Browsing WWW	R. Armstrong et al., "WebWatcher: A Learning Apprentice for the World Wide Web," AAAI 1995 Spring Symp. Information Gathering from Heterogeneous, Distributed Environments, AAAI Press, Menlo Park, Calif., 1995

Table 2 Agents using collaborative techniques

Agent	Organization	Goal	Publication
Firefly, Ringo	MIT	Finding music, movie, books	P. Maes, "Agents that Reduce Work and Information Overload," Comm. ACM, Vol. 37, No. 7, July 1994, pp. 30–40
GroupLense	Minnesota	Usenet news filtering	J.A. Konstan et al., "GroupLense: Applying Filtering to Usenet News," Comm. ACM, Vol. 40, No. 3, Mar. 1997, pp. 77–87
Phoaks	AT & T labs	Browsing WWW	T. Terveen et al., "PHOAKS: A System for Sharing Recommendations," Comm. ACM, Vol. 40, No. 3, Mar. 1997, pp. 59–62

5 Conclusion

Intelligent agents are becoming increasingly useful in search engines [28] and in indexing and spidering applications to help with information overload and filtering out information that is most relevant. They help with data mining and analysis to glean important information. Currently, searching is passive and static. In the near future, intelligent Agents could make search engines active personal assistants that would proactively weed out information irrelevant to users, using user preference algorithms and advanced information retrieval techniques, thereby making browsing more intuitive and tailored to user preferences.

Acknowledgments I extend my heartfelt gratitude to my mentor, Ms. Garima Srivastava, Asst. Professor, Dept of Computer Science and Engineering, Amity University, Lucknow for her guidance

and kind encouragement. I am thankful to my family and friends without whose support this paper would not have been possible.

References

1. Lalmas, M., van Rijsbergen, K.: A logical model of information retrieval based on situation theory. In: 14th Information Retrieval Colloquium. Springer, London (1993)
2. Koller, D., Sahami, M.: Hierarchically classifying documents using very few words. Stanford InfoLab (1997)
3. Anonymous.: The age of the intelligent agent. *Insur. Syst. Bull.* **9**(10), 4–5 (1994)
4. Mansour, K., Hattab, E., Abuelrub, E.: Intelligent Agents Review. In: iiWAS. (2004)
5. Kolodner, J.: Case-Based Reasoning (Morgan Kaufmann Series in Representation and Reasoning). Morgan Kauffman (1993)
6. Lenz, M., Hübner, A., Kunze, M.: Question answering with textual CBR. In: International Conference on Flexible Query Answering Systems. Springer, Berlin, Heidelberg (1998)
7. Russell, S., Norvig, P.: Artificial Intelligence. A modern approach. Artificial Intelligence. Prentice-Hall, Englewood Cliffs 25.27: 79–80 (1995)
8. UMBC AgentNews Web Letter, Agents on the net, vol. 1, No. 17, December, 1996
9. Lieberman, H.: Letizia: An agent that assists web browsing. *IJCAI* (1) 924–929 (1995)
10. Jaczynski, M., Trousse, B.: Broadway: a world wide web browsing advisor reusing past navigations from a group of users. In: Proceedings of the Third UK Case-Based Reasoning Workshop (UKCBR3) (1997)
11. Aiken, M., Govindarajulu, C.: Knowledge-based information retrieval for group decision support systems. *J. Database Manag.* **5**(1), 1–35 (1994)
12. Beheshti, J.: Browsing through public access catalogs. *Inf. Technol. Libr.* **11**(3), 220–228 (1992)
13. Jansen, J.: Using an intelligent agent to enhance search engineer performance. *Online* **2**(3) (1997)
14. Burke, R., Hammond, K., Kozlovsky, J.: Knowledge-based information retrieval from semi-structured text. Working Notes from AAAI Fall Symposium on AI Applications in Knowledge Navigation and Retrieval (1995)
15. Goldberg, D., Nichols, D., Oki, B.M., Terry, D.: Using collaborative filtering to weave an information tapestry. *Commun. ACM* **35**(12), 61–70 (1992)
16. Godoy, D., Analía, A.: Personal Searcher: an intelligent agent for searching web pages. In: Advances in Artificial Intelligence, pp. 43–52. Springer, Berlin, Heidelberg (2000)
17. Pazzani, M., Billsus, D.: Learning and revising user profiles: the identification of interesting web sites. *Mach. Learn.* **27**(3), 313–331 (1997)
18. Papadimitriou, D., Koutrika, G., Velegrakis, Y., Mylopoulos, J.: Finding related forum posts through content similarity over intention-based segmentation. *Knowl. Data Eng. IEEE Trans.* **29**(9), 1860–1873 (2017)
19. Kobayashi, M., Takeda, K.: Information retrieval on the web. *ACM Comput. Surv. (CSUR)* **32**(2), 144–173 (2000)
20. Lashkari, Y., Metral, M. Maes, P.: Collaborative interface agents. In: Proceedings of AAAI '94 Conference, Seattle, Washington, August 1994
21. Bhute, A.N., Bhute, H.A., Meshram, B.B.: Intelligent web agent for search engines. In: International Conference on Trends and Advances in Computation and Engineering, TRACE (2010)
22. Joachims, T., Freitag, D., Mitchell, T.: Webwatcher: a tour guide for the world wide web. *IJCAI* **1** (1997)
23. Mladenic, D.: Browsing Assistant Personal Webwatcher. Machine learning used by Personal WebWatcher (1999)
24. Mladenic, D.: Personal WebWatcher: Implementation and Design, Technical Report IJS-DP-7472, Carnegie Mellon University, Pittsburgh (1996)

25. Godoy, D., Analía, A.: PersonalSearcher: an intelligent agent for searching web pages. *Advanc. Artif. Intell.* Springer, Berlin, Heidelberg, pp. 43–52 (2000)
26. Sebastiani, Fabrizio: Machine learning in automated text categorization. *ACM Comput. Surv. (CSUR)* **34**(1), 1–47 (2002)
27. Smyth, B., Paul, C.: Surfing the digital wave. In: *International Conference on Case-Based Reasoning*. Springer, Berlin, Heidelberg (1999)
28. Schwartz, Candy: Web search engines. *J. Am. Soc. Informat. Sci.* **49**(11), 973–982 (1998)

Impact of Internet of Things on Societal Applications



Tanya Srivastava and Shikha Singh

Abstract The technology of Internet of things, also curtailed as IoT, is an emerging technology these days and is often used to connect various appliances or the gadgets that are generally used by us in our daily life with the facility of internet so as to make us capable of sending commands to these connected devices in order to perform a piece of work or an action in a desired manner without the involvement of any human labor. It is introduced to reduce the mankind labor that is generally involved in performing certain activities and to spawn everything in an automatic environment that works on the basis of commands. IoT technology has its major emphasis on the connection between the “things” and the Internet. This “thing” mentioned here can be a living thing, which may include a man or an animal, so as to sense the occurrence of various activities that take place in them. It can also be a non-living thing including the gadgets or any other electronic appliances. IoT incorporates various devices that are enabled by an active connection to the Internet and are able to transfer information collected by them. These devices then act upon the records collected from their surroundings through sensors that are rooted in them. IoT technology comes up with a large range of applications these days. Some of their uses are in smart watches, smart home security and household appliances, embedded cars, and many more. In the past few years, IoT has emerged out as an innovative technology by converting the normal conventional industries into smart industries by its numerous applications in the field of Industrial IoT and in our ordinary homes to smart homes by its applications in Home Automation System. This review paper has its emphasis on what is IoT, its working, its various applications in industrial purposes, its beneficial applications to humans and its affect in our daily lives. It also consists of a few setbacks IoT technology and various methods of removing those drawbacks so as to make it more beneficial to mankind.

T. Srivastava (✉) · S. Singh

Department of Computer Science and Engineering, Amity School of Engineering and Technology, Amity University, Lucknow Campus, Lucknow, Uttar Pradesh, India
e-mail: srivastavatanya12@gmail.com

S. Singh
e-mail: ssingh8@lko.amity.edu

Keywords Sensors • IOT technology • Internet

1 Introduction

The basic idea of implementing the Internet of Things (IoT) technology was primarily offered by Kevin Ashton. He is a British technology pioneer and one of the founders of the AutoID located at Massachusetts Institute of Technology. He was the first person to lay down the foundation of this in his presentation in the year 1999. On the basis of his research, in today's era, the human-generated computers and systems are totally dependent on humans for the sake of accessing information by the medium of Internet. When he was doing his research back in the year 1997, he concluded that approximately 50 petabytes of data that is available and accessible by the Internet was all human-created data. It is an acceptable fact that humans can invest very less time from their busy schedule to manage many tasks at a time, but have a desire to expect accuracy and perfection in all domains Based on this fact, Kevin realized that if the computers are fed with all the information about humans that it could know and about the belongings we frequently require, then these computers systems can be proved of great help to mankind. According to his idea, the systems should be designed to track important information from things we frequently require and with the help of this data collected, it should be capable of performing tasks which were basically supposed to be completed by humans. As a result, it will reduce human effort and can attain the efficiency which normal humans can't [1].

The concept behind IoT can be briefly summarized in the following way- "*It is a concept which describes the proposal of connecting the frequently used gadgets and appliances to the internet in order to implement autonomous functioning of the gadgets in accordance to what humans desire, each of these devices being capable of defining itself to other such gadgets or appliances [2].*"

2 How Internet of Things Work

The game-changing technology of IoT is also commonly referred to as Internet of Everything (IoE). It is because it holds the capability of performing almost everything which can be performed by humans. This is the reason why this technology is introduced, to reduce the efforts which humans have to make. IoT devices enabled by an active internet connection that are capable of sending or collecting information or basically transferring data and collected by the sensors embedded in them. The IoT enabled devices comprises of three basic components, namely **sensors, data processing units, and communication facilitating hardware components**. The use of sensors is collection of data from its environment, data processing units to process on that data, and finally communication facilitating hardware components give a response to those actions that are sensed by sensors. Such devices are commonly

known as “Smart devices”, because these devices are capable of communicating with other devices by the means of Machine to Machine Communication or M2M Communication, and perform actions in accordance to the collected data or information. Some famous applications of IoT are Smart homes, Wearable, cities, Industries, Connected cars, etc. [1]

The basic working of IoT primarily consists of RFID (Radio Frequency Identification) chips. These microchips facilitate the transfer of identification information to the devices by transmitting it via a wireless means of communication. By the means of RFID tags, we can be able to identify or track any item. WSN (wireless sensor networks) is also used, which by the means of interconnected sensors, senses and monitors the items [3, 6]. This technology also uses the interconnected sensors to perform sensing and monitoring actions. WSNs are involved in many applications such as monitoring environment, health, industries, traffic, and many more [6, 7]. With the advancement of IoT day by day, several other advanced technologies are being brought into application to extend the potential of IoT. Mentioning a few of these technologies, IoT also incorporates barcodes scanners, smart mobile phones, cloud computing, etc. [4, 5]. IoT has now become a huge part of industries in the past few years and because of the dipping cost of IoT enabled devices, companies are welcomed to implant smaller and frequently used objects with RFID chips [8].

3 Service Oriented Architecture of IoT

Since IoT technology is a key behind the connecting varied type of devices, the proposal of service-oriented architecture (SOA) is appropriate for it. It is a successful method of organizing components in an architecture in the fields of cloud computing and WSNs. It consists of a multilayer SOA structure on selected technologies and technical requirements. A Service Oriented Architecture consists of four layers and a brief introduction is mentioned below. This architecture of SOA is briefly described in Table 1.

Table 1 Four layers of Service-Oriented Architecture (SOA)

Layers	Description
Sensing layer	This layer of SOA is integrated with RFID, sensors, etc. to sense and control the physical things in the world and receive data from them
Networking layer	This layer of SOA provides the basic networking support to the system and transfer of data over a network, wired or wireless
Service layer	This layer of SOA is used to create and manage services to satisfy the needs of the user
Interface layer	This layer of SOA provides methods of interaction to the users

Since IoT requires the capability of efficiency in execution of tasks, it is concluded that SOA approach forms a good architecture to accomplish the feature of interoperability among the devices [9–11].

3.1 Sensing Layer

The sensing layer enables the wireless smart devices and gadgets to sense involuntarily and share the sensed information among devices. A UUID or universal unique identifier is fixed in these devices to enhance identification process [12, 13].

3.2 Networking Layer

The networking layer plays the role of connecting all the IoT enabled devices collectively and allow them to transfer and share the information among all the connected devices [14].

3.3 Service Layer

The service layer architecture is dependent upon middleware expertise in which implicational uses are provided to assimilate the services and the applications of IoT, which are responsible for the provision of a cost-efficient proposal, to facilitate the reusability of the software as well as the hardware platforms [9, 10, 14].

3.4 Interface Layer

This layer is responsible for simplifying the problems that are related to supervision and interconnection of gadgets. An interface profile or IFP is used as a standard of the interface layer in order to support interface with various applications. It is associated to the performance of UPnP protocol, or Universal Plug and Play protocol for facilitating the dealings and relations of these gadgets with services offered by them [14, 15].

4 Industrial Applications of Internet of Things

Even though the IoT is still emerging, its applications are growing day by day. Some of these applications IoT prevailing in the industries are described as follows:

- Uses of IoT in biological health industry [16]. Using the identification, sensing, and communication capabilities, all the objects of hospitals (people, equipments, medicine, etc.), can easily be kept track of and monitored [17]. Consider the example of a heart patient, his heart rate can be sensed by the means of sensors and the data can be directly forwarded to the physician from time to time involving the patient. By using personal devices and internet, various services provided by IoT can be personalized [18].
- IoT contribution in the industry of food [19]. The food supply chains, also called FSCs, are exceedingly complex chains scattered all over the world. But its complexity has as an obstruction in the supervision of its efficiency. Its contribution ranges from strict and clear-cut agriculture techniques to productivity of food, dispensation, storage, and delivery. It comprises three basic components, the field like WSN and RFID, databases and servers and the final part, the communication infrastructures such as WLAN, satellite, Ethernet, etc.
- IoT in transport and shipment [9]. A lot of objects these days are prepared within built bar codes and RFID chips. This enables the industries of transportation and shipment to carry out real-time monitoring of the shipped [20]. The development of an autopilot application is expected to be seen in future which can be capable of detecting the pedestrians and other vehicles and handles the steering of the vehicle robotically to avoid accidents [21].
- Use IoT in firefighting techniques. RFID tags are attached to the equipments of fighting with fire in many countries, by means of which a database is generated to manage such. With the assistance of these RFID tags, readers, cameras, sensors, and various wireless mediums of communication, the firefighting departments could perform environmental monitoring in real time as soon as possible and start the rescue operation as per requirement.

5 Internet of Things in Smart Cities

In a research, it was predicted that about 70% of the world's population will be residents of urbanized cities by the year 2050. This speedy expansion of urban inhabitants is producing a lot of pull on the infrastructural companies to provide accommodation to the people relocating to urban areas [22].

Here are a few applications proposed by IoT for smart cities.

5.1 Efficient Delivery of Water

The prior methods of water delivery have been transformed by the intervention of IoT in the cities. The introduction of smart meter usage has improved the exposure of leakage in the water tank is one of its most beneficial application, making the water supply system more proficient.

5.2 Traffic Blocking Solution

Traffic signals integrated with IoT, are capable of adjusting the timing of signals by sensing the traffic congestion. A cumulative data about traffic conditions can be collected from the cameras, and sensors on roads for monitoring traffic in real-time.

5.3 Public Transportation Facilities

IoT helps in this case by providing real-time warnings that are required to enforce the implementation of plans to avoid interrupts top people, providing a safe and efficient public transportation facility. This can be by the use of cameras or connected bus stops or other public areas.

5.4 Saving Energy in Buildings

IoT incorporates the use of energy management systems, making connections to lighting and fire-safety equipments to a central management system. It also helps in highlighting the areas of high electricity usage and rectify the faults that occur at those areas.

5.5 Public Safety Improvement

A video analyzing software is used to processes the video footage recorded by the cameras for spotting only the important events, ignoring the unnecessary information. The security systems use IoT to turn every attached camera into the sensor, along with end-to-end computing and analyzing from the source.

6 Internet of Things in Smart Cars

From the viewpoint of linked vehicles, a few of the major appropriate ideas powered by IoT technology which may soon be enforced are-

- Response to vehicle crash: Connected cars these days are proficient of automatically transferring real-time information about the incident along with its location to emergency centers.
- Diagnosis of Car Troubles: Connected cars are capable of generating data that can be used for prediction of any upcoming failure before a component fails, preventing the cause of inconvenience.
- Convenient Services: It provides the ability to spot a car remotely, making it easy to access the services like far-off door unlocking and find my vehicle.
- Navigation: It enables the connected cars to be integrated with GPS system which responds to direction-finding, fuel alerts, traffic, etc.
- Traffic Management: It provides the transportation agencies with real-time traffic information and parking data to make it easier for them to manage.

7 Pros and Cons of Industrial Internet of Things

Other than emerging at a fast rate, IoT has also become so famous and well known to every person that it no more requires to be introduced. When the applications of services of IoT are deployed in industrial and manufacturing areas, it is often referred to as **Industrial IoT**. Some of its pros and cons are mentioned below [22].

7.1 *Pros of Industrial Internet of Things*

- Communication systems: IoT has been encouraging the communication among the gadgets and appliances in numerous well-situated ways, known as Machine-to-Machine (M2M) communication which makes the devices proficient in remaining connected with other devices.
- Autonomous behavior and Control: There prevails a large amount of automatic work in the manufacturing industries. The machines could now communicate on their own, resulting in faster and accurate output.
- Optimal use of time and funds: It is considered as the major advantage of IoT. It promotes saving of precious time and money. Even the charge of monitoring the IoT equipment has fallen down to a great extent.
- Efficiency: In addition to all the above advantages, it also helps in escalating the efficiency of the system. The reason is that the machine-to-machine interaction provides an efficient system with rapid and precise consequences.

7.2 Cons of Industrial IoT

- Compatibility with other devices: IoT has no international compatibility standard for monitoring the devices and equipments.
- Security and privacy: IoT connects the devices to Internet which provides the availability of all data on the Internet making it prone to leakage.
- Unemployment: Because of the introduction of automatic working in everything, it has caused the loss of jobs of many workers and laborers.

8 Conclusion

Viewing the complexity of the IoT technology, it connects the devices together that are equipped with sensors for identification, data processing units, communication, and wireless networking capabilities. With the progression of the technology, sensors being used are becoming progressively powerful in the course of time with reduction in its cost dimensions, making it an exceedingly pleasing invention in the marketplace. A number of research works are going on in the industry to deploy these products for industrial expansion applications and undoubtedly, it is proved to be among the most widely functional inventions in the industries.

This paper presents a review work of several research works on Industrial IoT. It begins with introducing the background of IoT and service-oriented architectural approach, and various applications offered by it. There is also a discussion about various pros and cons of IoT. The main involvement of this paper is to focus on industrial IoT and its applications.

References

1. Johnson B.: How the Internet of Things works
2. Techopedia: Internet of Things (IOT)
3. Jia, X., Feng, O., Fan, T., Lei Q.: RFID technology and its applications in Internet of Things (IoT). In: Proceedings of the 2nd IEEE International Conference on Consumer Electronics, Communications and Networks (CECNet). Elissa, K. "Title of paper if known," unpublished, pp. 1282–1285, April 21–23 (2012)
4. Uckelmann, D., Harrison, M., Michahelles, F.: An architectural approach towards the future internet of things. In: Architecting the Internet of Things, pp. 1–24. Springer (2011)
5. Li, Q., Wang, Z., Li, W., Li, J., Wang, C., Du, R.: Applications integration in a hybrid cloud computing environment: modelling and platform. Enterp. Inf. Syst. **7**(3), 237–271 (2013)
6. Li, S., Xu, L., Wang, X.: Compressed sensing signal and data acquisition in wireless sensor networks and Internet of Things. IEEE Trans. Industr. Inf. **9**(4), 2177–2186 (2013)
7. He, W., Xu, L.: Integration of distributed enterprise applications: a survey. IEEE Trans. Industr. Inf. **10**(1), 35–42 (2014)
8. Thrasher, J.: A Primer on the Internet of Things & RFID. Atlasrfidstore (2014)
9. Atzori, L., Iera, A., Morabito, G.: The internet of things: a survey. Comput. Netw. **54**(15), 2787–2805 (2010)

10. Miorandi, D., Sicari, S., De Pellegrini, F., Chlamtac, I.: Internet of things: vision, applications and research challenges. *Ad Hoc Netw.* **10**(7), 1497–1516 (2012)
11. Xu, L.: Enterprise systems: state-of-the-art and future trends. *IEEE Trans. Ind. Inform.* **7**(4), 630–640 (2011)
12. Wu, Y., Sheng, Q.Z., Zeadally, S.: RFID: opportunities and challenges. In: Chilamkurti, N. (ed.) *Next-Generation Wireless Technologies*, pp. 105–129, Springer (2013)
13. Ilie-Zudor, E., Kemeny, Z., Van Blommestein, F., Monostori, L., Van Der Meulen, A.: A survey of applications and requirements of unique identification systems and RFID techniques. *Comput. Ind.* **62**(3), 227–252 (2011)
14. Guinard, D., Trifa, V., Karnouskos, S., Spiess, P., Savio, D.: Interacting with the soa-based internet of things: discovery, query, selection, and on-demand provisioning of web services. *IEEE Trans. Serv. Comput.* **3**(3), 223–235 (2010)
15. Gama, K., Touseau, L., Donsez, D.: Combining heterogeneous service technologies for building an Internet of Things middleware. *Comput. Commun.* **35**(4), 405–417 (2012)
16. Pang, Z., Chen, Q., Tian, J., Zheng, L., Dubrova, E.: Ecosystem analysis in the design of open platform-based in-home healthcare terminals towards the internet-of-things. In: Proceedings of 2013 15th International Conference on Advanced Communication Technology (ICACT), pp. 529–534. Jan 27–30 (2013)
17. Alemdar, H., Ersoy, C.: Wireless sensor networks for healthcare: a survey. *Comput. Netw.* **54**(15), 2688–2710 (2010)
18. Plaza, I., MartíN, L., Martin, S., Medrano, C.: Mobile applications in an aging society: status and trends. *J. Syst. Softw.* **84**(11), 1977–1988 (2011)
19. Pang, Z., Chen, Q., Han, W., Zheng, L.: Value-centric design of the internet-of-things solution for food supply chain: value creation, sensor portfolio and information fusion. *Inf. Syst. Front.* (2014) (in press)
20. Karakostas, B.: A DNS architecture for the Internet of Things: a case study in transport logistics. *Procedia Comput. Sci.* **19**, 594–601 (2013)
21. Keller, C.G., Dang, T., Fritz, H., Joos, A., Rabe, C., Gavrila, D.M.: Active pedestrian safety by automatic braking and evasive steering. *IEEE Trans. Intell. Transp.* **12**(4), 1292–1304 (2011)
22. Sannapureddy, B.R.: Pros and Cons of Internet of Things(IOT), Feb 25 (2015)

Experimental Validation of Fully Informed Particle Swarm Optimization Tuned Multi-Loop L-PID Controllers for Stabilization of Gantry Crane System



Sudarshan K. Valluru, Madhusudan Singh, Daksh Dobhal,
Kumar Kartikeya, Manpreet Kaur and Arnav Goel

Abstract Linear PID controllers are commonly used as an electrical component to decrease the error between anticipated set value and actual measured values for control of various benchmarked systems. The Multi-Loop Linear PID (ML-PID) controller, gives a robust and efficient performance in most of the situations. This paper presents the implementation of linear PID controller to stabilize and control the Gantry Crane System. Optimal performance is obtained for a few specific combinations of the proportional, integral, and derivative gains, which makes it essential to tune these values through Optimization techniques. The Fully Informed Particle Swarm Optimizer (FIPSO) is used to tune the gain values of the ML-PID controlled Gantry Crane System. These values are validated experimentally, and obtained results prove that the FIPSO tuned multi-loop linear PID controller quickly stabilizes the system subjected to external disturbances.

Keywords Multi-Loop linear PID (ML-PID) · Fully Informed Particle Swarm Optimization (FIPSO) · Gantry crane system

S. K. Valluru (✉) · M. Singh · D. Dobhal · K. Kartikeya · M. Kaur · A. Goel

Department of Electrical Engineering, Incubation Center for Control Dynamical Systems and Computation, Delhi Technological University, New Delhi 110042, Delhi, India

e-mail: sudarshan_valluru@dce.ac.in

M. Singh
e-mail: madhusudan@dce.ac.in

D. Dobhal
e-mail: dakshdobhal@gmail.com

K. Kartikeya
e-mail: kkartikeya1998@gmail.com

M. Kaur
e-mail: manpreetdtu.26@gmail.com

A. Goel
e-mail: arnav_bt2k16@dtu.ac.in

1 Introduction

The Linear PID is one of the most widely used electrical component in the systems industry [1, 5]. It provides efficient solutions while being computationally inexpensive and simple to implement and simulate. The Linear PID can control and modulate responses of benchmarked dynamical systems [11, 12]. Though, if the system has nonlinearities, the controller may produce unsatisfactory response including high peak overshoot and large settling time, but still, it has limitations like optimality and tuning [10]. Thus, the difficulty of tuning is to be challenging particularly if the multi PID controllers have more than three objectives. The three terms in the controller namely Proportional, Integral, and Derivative can be varied to get an optimal response which is calculated according to the external disturbances by evolutionary computation [8].

In recent years, Particle Swarm Optimization has gained much attention of researchers as a bio-inspired algorithm for the optimization techniques especially for the multimodal and multi-output functions [6, 13]. Although the standard PSO showed better results and faster convergence time to reach global optima for unimodal objective functions, it suffers from falling of local minima problem for nonlinear, multimodal objective functions [6]. This paper addresses the simultaneous tuning of ML-PID controllers using a new variant of FIPSO [2, 4], and the obtained gain values are tested to control and stabilize the Gantry Crane System [3].

The rest of the paper is organized as following sections. Section 2 describes the design of a Multi-Loop Linear PID Controller. Mathematical modeling of the Gantry Crane system is presented in Sect. 3. The tuning of the Controller using FIPSO is presented in Sect. 4. Experimental results and discussions are given in Sect. 5, and finally, Sect. 6 outlines the conclusion of the paper.

2 Design of Multi-Loop L-PID Controllers

The linear proportional integral derivative controller operates upon the error term $\varepsilon(t)$, the mismatch between the real time output and reference input signals to generate the correction signal $\alpha(t)$. The correction signal is obtained as the sum of the error, derivative of the error, and integral of the error multiplied by their respective coefficients. Usually, the three coefficients are denoted by P, I, and D respectively. Hence the expression for the correction signal is obtained as (1):

$$\alpha(t) = K_p \varepsilon(t) + K_i \int_{t_1}^{t_2} \varepsilon(t) \cdot dt + K_d \frac{d\varepsilon(t)}{dt} \quad (1)$$

The primary task is determining the optimal combinations of the coefficients which yield the best performance. The term proportional to the error reduces the system response time as it is increased till a maximum allowed value above which

sustained oscillations are introduced in the system response. In reducing the steady state error, the major part is played by the integral term. The derivative term leads to a reduction of the corrected signal if the error is increasing.

3 Gantry Crane System

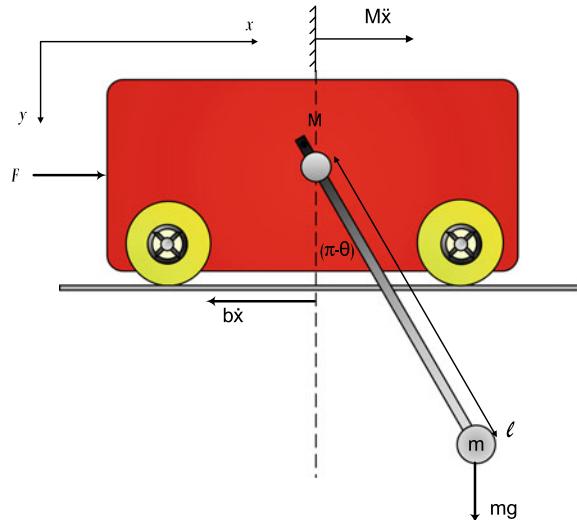
The crane system, in spite of its simple structure, is a complex system owing to its nonlinear nature [7]. Because of this, it is considered to be one of the under accutated benchmark of control systems. If it is assumed that the payload is a lumped-mass distribution rather than a distributed-mass system, the gantry crane can be modeled as a simple pendulum. Utilizing the assumptions of payload being a point mass distribution, massless cable, the cable of fixed length and not flexible and neglecting the damping coefficient of the cable, the free body diagram of the crane and functional diagrams are shown in Fig. 1 and Fig. 2 respectively. Consider M , m , l , b , and I to be the mass of the trolley, mass of the pendulum bob, length of the cable, the coefficient of friction of cart, and the MOI of the bob respectively.

The variables x and θ are the position of trolley along the belt and angular position of the pendulum. The equations of motion obtained for the system are (2) and (3):

$$F = (M + m) \cdot \dot{x} + b\dot{x} - mlsin\theta \cdot \dot{\theta}^2 + mlcos\theta \cdot \dot{\theta} \quad (2)$$

$$0 = (I + ml^2) \cdot \ddot{\theta} - mglsin\theta + mlcos\theta \cdot \dot{x} \quad (3)$$

Fig. 1 FBD of gantry crane system



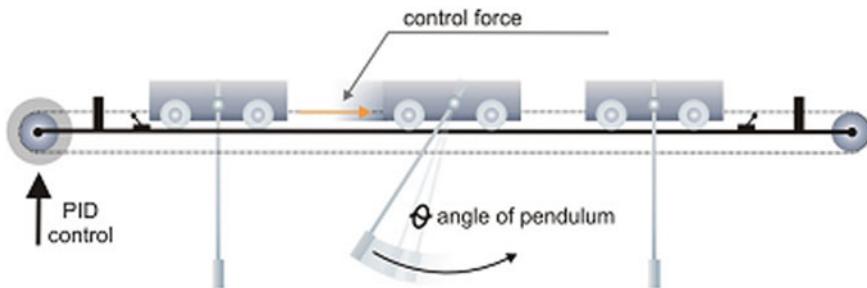


Fig. 2 Functional Diagram

Table 1 Gantry crane system parameters

S. No.	Parameters	Value
1	Mass of the cart (M)	2.4 kg
2	Mass of the pole (m)	0.23 kg
3	Cart friction coefficient (b)	0.05 Ns/m
4	Length of the pole (l)	0.32 m
5	Pendulum moment of inertia (I)	0.099 kg m ²
6	Gravity constant (g)	9.81 ms ⁻²
7	Range of operating voltage of D.C. servo motor (V)	±24 V

For $\theta \rightarrow 0$: $\sin\theta \approx \theta$, $\cos\theta \approx 1$, $\dot{\theta}^2 \approx 0$

For the simpler understanding of carrying out the stability analysis of the dynamics of the system, the linearized system Eqs. (4) and (5) as

$$F = (M + m) \cdot \ddot{x} + b\dot{x} + ml \cdot \ddot{\theta} \quad (4)$$

$$0 = (I + ml^2) \cdot \ddot{\theta} - mgl\dot{\theta} + ml \cdot \ddot{x} \quad (5)$$

The system parameters as specified in Eqs. (4) and (5) are given in Table 1.

4 Tuning of ML-PID by FIPSO Algorithm

The PSO algorithm mimics the flocking of birds and the schooling of fishes to find the optimal value of the given target function [9]. The Fully Informed Particle Swarm Optimization takes into account the influence on each particle in a swarm by the mean of its nearest neighbor. Therefore, it has the advantage of fast convergence, escaping

the tendency to fall in local minima and ability to optimize complex functions when compared to Standard PSO algorithm.

The important variables associated with particles in the swarm in FIPSO are x_i (position), v_i (instant velocity), Pb_i (Present Personal best), and Gb_i (Global Position best). Each particle is updated on the basis of above parameters in the above-defined pop size (iteration) for which cost Function (F_{cost}) is to be minimized. For the Gantry Crane Benchmark, x_i can be comprehended as (6)

$$x_i = [K_{P1}, K_{I1}, K_{D1}, K_{P2}, K_{I2}, K_{D2}] \quad (6)$$

where x is the position of the particle, K_{P1} , K_{I1} , K_{D1} are PID constants for the cart's position controller and K_{P2} , K_{I2} , K_{D2} are for the angle controller of the gantry crane system. Particles in a swarm were initialized using the Eq. (7)

$$x_i = unifrnd(PosLL_i, PosUL_i, npop, 1) \quad (7)$$

where *unifrnd* function generates uniform random numbers in the range ($PosLL_j, PosUL_j$). $PosLL$ And $PosUL$ are the lower and upper decision boundary variables respectively, $npop$ is the population (swarm) size and j th represents one of the PID constants shown in above equation.

The updated velocity and position of each particle in an iteration can be calculated as in Eqs. (8) and (9)

$$v_{i+1} = w * v_i + c_1 * (Pb_i - x_i) + c_2 * (Gb_i - x_i) + chi * (x_i - NN_i) \quad (8)$$

$$x_{i+1} = x_i + v_{i+1} \quad (9)$$

where NN_i is the Nearest Neighbor, defined as the mean of the personal best position of the nearest five particles w is the inertial weight, c_1, c_2 are personal and global learning coefficients respectively and chi is the constriction coefficient. Finally, the innovative Global Best and Personal Best are updated as (10) and (11):

$$Pb_{i+1} = \begin{cases} Pb_i, & Fcost(x_i) > Fcost(Pb_i) \\ x_i, & otherwise \end{cases} \quad (10)$$

$$Gb_{i+1} = \min\{Fcost(Pb_{i+1})\} \quad (11)$$

4.1 Implementation of FIPSO for ML-PID Tuning

The overall FIPSO pseudo code to tune the ML-PID controller for gantry system is summarized as follows:

```

FCost- call the function whose variables are to be
tuned
nVar- no. of variables to be tuned
NoVar- row vector defining size of the decision
variable matrix
PosLL- minimum values of decision variables
PosUL- maximum value of decision variables
VelUL-maximum value of velocity
VelLL- minimum value of velocity
Setting required PSO parameters:
Gen- maximum no of iterations to be undergone by the
code
PopSize- size of the swarm
Km- No. of nearest neighbours considered
NN- Best position closest of neighbours
Initialisation of swarm parameters
global best- array of size NoVar containing zeros
for i=1: PopSize
    initialize position of particle assigning it random
    value;
    set velocity of particle zero;
    evaluate cost function for particle;
    set personal best position and cost equal to
    initial position and cost for each particle;
    if personal best cost of particle i > global best
    cost
        global best=personal best of i;
    end
end

Main Loop of PSO
for it=1: Gen
    for i=1: PopSize
        update velocity using FIPSO innovative vector
        formula;
        apply velocity limits so that velocity remains
        between VelLL and VelUL;
        update position;
        apply position limits so that position remains
        between PosLL and PosUL;
        calculate value of cost function;
        if cost of i < personal best cost of i
            personal best position of i= position of i;
            personal best cost of i= cost of i;
        end
        if global best cost < best cost of i
            global best= personal best of i;
        end
    end
end

apply position limits so that position remains between
VarMin and VarMax;
calculate value of cost function;
if cost of i < personal best cost of i
personal best position of i= position of i;
personal best cost of i= cost of i;
end
if global best cost < best cost of i
global best= personal best of i;
end
end
end

```

Table 2 ML-PID gain parameters

S. No	Gain co-efficient	Optimal value
1	K_{P1}	4.8590
2	K_{I1}	0.4918
3	K_{D1}	4.0352
4	K_{P2}	4.9101
5	K_{I2}	0.0973
6	K_{D2}	0.3000

The swarm size, iteration, w , c_1 , c_2 , chi values used in the experiment for the optimization of ML-PID controller for the gantry crane system are 50, 30, 0.7298, 1.4962, 1.4962, 0.7298 respectively. Two main objectives for control of Gantry Crane System, namely, swing minimization of pendulum and tracking of input voltage signal, were achieved using two separate PID controllers. The optimal gain values obtained at 30th iteration with minimum cost function is given in Table 2.

5 Experimental Results

The cart was required to track a sinusoidal trajectory of amplitude 0.3 V and a frequency of 0.1 Hz. A sample time of 100 s was chosen to operate the system. Small manual perturbations of increasing magnitude were given to pendulum for every 20s intervals. Cart Position Response, Pendulum Angle Response, and Control Signal depicting the performance of ML-PID were obtained and validated extensively for stabilizing the Gantry Crane System as shown in Fig. 3. The measure of performance

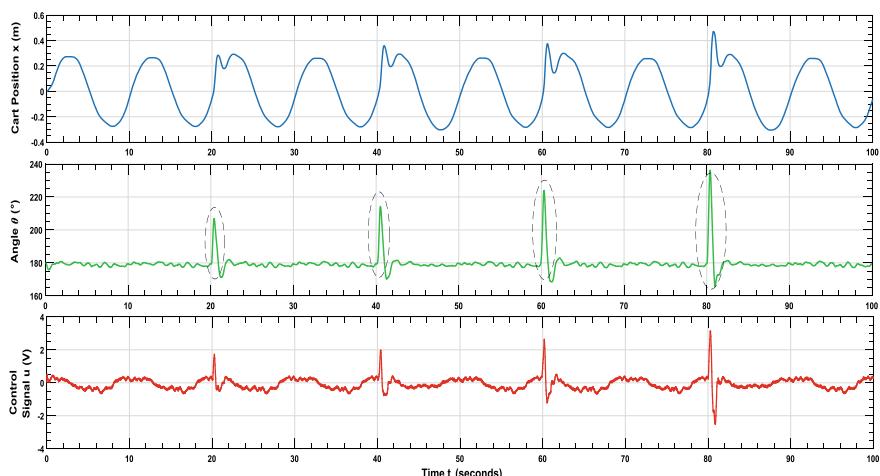


Fig. 3 Performance of Multi-Loop L-PID controller

Table 3 Pendulum angle stabilization response

S. No.	Instant of disturbance initiation (s)	Instant of stabilization (s)	Amplitude of disturbance ($^{\circ}$)	Time taken to stabilize (s)
1	20.11	21.57	26.9	1.46
2	40.4	41.7	34.3	1.3
3	59.96	61.58	43.9	1.72
4	80.06	81.67	56.3	1.61

was time taken for stabilization after an external disturbance was given. The instant of stabilization was determined according to the 3% settling time stability condition.

The ML-PID responds to stabilization of the pendulum angle against the external noise/disturbance is given in Table 3. It is observed that as the magnitude of disturbance increases, the time taken for stabilization also increases.

6 Conclusions

In this paper, ML-PID controllers were designed, and the optimal gain parameters are obtained by using the FIPSO algorithm. The controllers were validated to stabilize the benchmarked gantry crane system. It is observed that the ML-PID controller brings the pendulum to the desired position under the perturbations in an average time of 1.522 s.

Acknowledgements The authors would like to thank the Govt. of India and Govt. NCT Delhi for providing funding through Delhi Technological University under TEQIP-II to procure Digital Pendulum System Experimental Setup.

References

1. Astrom, K.J., Hagglund, T.: Benchmark systems for PID control. In: IFAC Proceedings on Digital Control: Past, Present and Future of PID Control, pp. 165–166 (2000)
2. Gao, Y., et al.: Multi-objective opposition-based learning fully informed particle swarm optimizer with favour ranking. In: Proceedings—2013 IEEE International Conference on Granular Computing, GrC 2013, pp. 114–119 (2013)
3. Jaafar, H.I., Mohamed, Z., Abidin, A.F.Z., Ab Ghani, Z.: PSO-tuned PID controller for a non-linear gantry crane system. In: IEEE International Conference on Control System, Computing and Engineering, pp. 23–25, Penang, Malaysia (2012)
4. Mendes, R., et al.: The fully informed particle swarm: simpler, maybe better. IEEE Trans. Evol. Comput. **8**(3), 204–210 (2004)
5. Passino, K.M.: Intelligent control for autonomous systems. IEEE Spectr. **32**(6), 55–62 (1995)
6. Poli, R., et al.: Particle swarm optimization an overview. Swarm Intell. **1**(1), 33–57 (2007)

7. Ramli, L., et al.: Control strategies for crane systems: a comprehensive review. *Mech. Syst. Signal Process.* **95**, 1–23 (2017)
8. Reynoso Meza, G., et al.: Controller tuning by means of multi-objective optimization algorithms: a global tuning framework. *IEEE Trans. Control Syst. Technol.* **21**(2), 445–458 (2013)
9. Valluru, S.K., et al.: Comparative analysis of PID, NARMA L-2 and PSO tuned PID controllers for nonlinear dynamical system. *J. Autom. Syst. Eng.* **9**(2), 94–108 (2015)
10. Valluru, S.K.: *Introduction to Neural Networks, Fuzzy Logic & Genetic Algorithms*. Jaico Publishing House, Mumbai (2011)
11. Valluru, S.K., Singh, M.: Metaheuristic tuning of linear and nonlinear PID controllers to nonlinear mass spring damper system. *Int. J. Appl. Eng. Res.* **12**(10), 2320–2328 (2017)
12. Valluru, S.K., Singh, M.: Stabilization of nonlinear inverted pendulum system using MOGA and APSO tuned nonlinear PID controller. *Cogent Eng.* **4**(1), 1–15 (2017)
13. Zhao, J., et al.: Application of particle swarm optimization algorithm on robust PID controller tuning. In: Wang, L., Chen, K., Ong, Y.S. (eds.) *Advances in Natural Computation*, pp. 948–957. Springer, Berlin, Heidelberg (2005)

Design and Experimental Implementation of Multi-loop LQR, PID, and LQG Controllers for the Trajectory Tracking Control of Twin Rotor MIMO System



Sudarshan K. Valluru, Madhusudan Singh, Ayush and Abhiraj Dharavath

Abstract One DOF (degree of freedom) Twin Rotor multi-input multi-output system (TRMS) is one of the benchmarked dynamical systems which has striking comparability with Helicopter. Design of controllers for such complex air vehicle is a daunting task as it has inherent nonlinearities and cross-coupling between its main and tail rotors. Therefore, it is an interesting laboratory benchmarked control problem, to test the performance of various controllers. The control techniques focused in this paper are the experimental implementation of multi-loop controllers such as Linear Quadratic Regulator (LQR), linear proportional integral derivative (L-PID), and Linear Quadratic Gaussian (LQG) controllers on TRMS system to verify the performance of these controllers.

Keywords Twin rotor MIMO system · Linear quadratic regulator (LQR) · PID · Linear quadratic Gaussian (LQG)

1 Introduction

Nowadays there has been a remarkable interest in aerodynamical systems, the TRMS is one such system which has striking similarity like a helicopter, but it does not fly [3]. The Helicopter is highly nonlinear with unstable system dynamics during the operating regime [2, 6]. The main and tail rotors provide horizontal and vertical movement

S. K. Valluru (✉) · M. Singh · Ayush · A. Dharavath

Department of Electrical Engineering, Incubation Center for Control Dynamical Systems and Computation, Delhi Technological University, New Delhi, Delhi 110042, India
e-mail: sudarshan_valluru@dce.ac.in

M. Singh
e-mail: madhusudan@dce.ac.in

Ayush
e-mail: ayushaero@gmail.com

A. Dharavath
e-mail: dharavathabhiraj@gmail.com

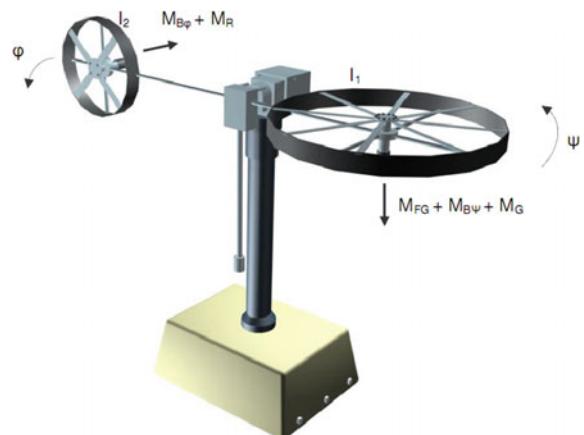
due to which it is to make land and take off vertically. There are two rotors situated at the ends of the beam which are main and tail rotors actuated by two DC servo motors. The main rotor changes the pitch angle whereas the tail rotor changes the yaw angle. At the pivot, the beam is counterbalanced with a certain weight to settle the bar. The state of the beam is expressed with four control variables such as horizontal angle, vertical angle and their respective angular velocities which are measured by encoders and tachogenerators. However, the TRMS aerodynamic force is controlled by simultaneous speed variation of DC servo motors where the voltage is the control input. The task of these controllers will be to change the DC servo motor voltages depend on control variables. An experimental set up of TRMS is considered for the design of LQR, L-PID, and LQG controllers [4, 5]. The real-time implementation is done by employing these controllers in horizontal and vertical control loops to achieve the desired system response and performance. Comparative analysis is made between these controllers in terms of rise time, settling time, overshoot and steady state error. The layout of this paper is arranged in five sections commencing with an introduction followed by Sect. 2, which presents the mathematical description of the system, and also the linearization method is discussed. In Sect. 3, design methodologies of LQR, L-PID, and LQG are elucidated with the determination of parameters for experimental implementation. Section 4 presents the obtained results for three design methods, followed by a conclusion in Sect. 5.

2 Benchmarked Mathematical Model of TRMS

The phenomenological model of TRMS [1] is shown in Fig. 1 which are nonlinear that implies not less than one of the states (pitch angle ψ and its derivative or yaw angle φ and its derivative) is an argument of the nonlinear function.

$$I_1 \ddot{\psi} = M_1 - M_{FG} - M_{B\psi} - M_G \quad (1)$$

Fig. 1 TRMS phenomenological model



where $M_1 = a_1 \cdot \tau_1^2 + b_1 \tau_1$, $M_{FG} = M_g \cdot \sin \psi$, $M_{B\psi} = B_{1\psi} \cdot \dot{\psi} + B_{2\psi} \cdot \text{sign}(\dot{\psi})$,

$$M_G = K_{gy} \cdot M_1 \cdot \dot{\varphi} \cdot \cos \psi$$

where, M_1 is nonlinear characteristics, M_{FG} is gravity momentum, $M_{B\psi}$ is frictional force momentum, M_G is gyroscopic momentum. The main motor momentum is approximated by the transfer function of the DC servo motor and control circuit in s-domain is given in Eq. (2)

$$T_1 = \frac{k_1}{T_{11} + T_{10}} \cdot u_1 \quad (2)$$

The Eqs. (3), (4), and (15) are represented for horizontal plane motion

$$I_2 \cdot \ddot{\varphi} = M_2 - M_{B\varphi} - M_R \quad (3)$$

$$M_2 = a_2 \cdot \tau_2^2 + b_2 \cdot \tau_2 \quad (4)$$

$$M_{B\varphi} = B_{1\varphi} \cdot \dot{\psi} + B_{2\varphi} \cdot \text{sign}(\dot{\psi}) \quad (5)$$

The cross-coupling momentum M_R is given by Eq. (6)

$$M_R = \frac{k_c(T_0 s + 1)}{T_p s + 1} \cdot \tau_1 \quad (6)$$

The tail rotor momentum of the DC servo motor with the control circuit is given in Eq. (7)

$$\tau_2 = \frac{k_2}{T_{21}s + T_{20}} \cdot u_2 \quad (7)$$

The TRMS model parameters used in the above equations are listed in Table 1.

The TRMS can be linearized across equilibrium points X_0 , where $X_0 = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$, Where X and Y are state and output vector given in Eqs. (8) and (9).

$$X = [\Psi \ \dot{\Psi} \ \varphi \ \dot{\varphi} \ \tau_1 \ \tau_2 \ M_R]^T \quad (8)$$

$$Y = [\psi, \varphi]^T \quad (9)$$

The state space equations can be expressed as (10) and (11), respectively.

$$\dot{x} = Ax + Bu \quad (10)$$

$$y = Cx \quad (11)$$

Table 1 TRMS parameters and values [1]

Parameters	Values
Main rotor moment of inertia (I_1)	6.8×10^{-2} (kg m ²)
Tail rotor moment of inertia (I_2)	2×10^{-2} (kg m ²)
Main rotor static characteristic parameter (a_1)	0.0135
Main rotor static constant (b_1)	0.0924
Tail rotor static characteristic parameter (a_2)	0.02
Tail rotor static constant (b_2)	0.09
Gravitational momentum (M_g)	0.32(N m)
Friction momentum for main rotor pitch($B_{1\psi}$)	6×10^{-3} (N m s/rad)
Friction momentum for tail rotor pitch ($B_{2\psi}$)	1×10^{-3} (N m s ² /rad)
Friction momentum for main rotor yaw ($B_{1\varphi}$)	1×10^{-1} (N m s/rad)
Friction momentum for tail rotor yaw ($B_{2\varphi}$)	1×10^{-2} (N m s ² /rad)
Gyroscopic momentum (K_{gy})	0.05 (s/rad)
Gain of the main motor (K_1)	1.10
Gain of the Tail motor (K_2)	0.80
Main motor denominator constant (T_{11})	1.10
Main motor denominator constant (T_{10})	1.0
Tail motor denominator constant (T_{21})	1.0
Tail motor denominator constant (T_{20})	1.0
Cross-coupling momentum parameter (T_P)	2.0
Cross-coupling momentum parameter (T_o)	3.50
Cross-coupling momentum gain (K_c)	-0.20

where

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -0.909 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0.21818 & 0 & -0.5 & 0 & 0 \\ -4.70588 & 0 & 1.3588 & 0 & 0 & -0.088235 & 0 \\ 0 & 0 & 0 & 4.5 & -50 & -5 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0.8 \\ -0.35 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

3 Design of LQR, PID and LQG Controller

The optimal controllers LQR linear PID and LQG controller for lateral and longitudinal movement of the TRMS system is designed and tested in real-time environment.

3.1 Linear Quadratic Regulator (LQR) Controller Design

The optimal control is concerned with the operation of the TRMS at minimum cost. Linear Quadratic problems are those in which the framework elements are given by straight differential equations and cost is expressed in terms of quadratic functions. The objective function is time integral and the sum of the control energy and transient energy. Counting the transient energy in the objective function, with the acceptable value of settling time and maximum overshoot. Here the control output of TRMS is obtained by varying the gain parameter (K) of the linear quadratic regulator. Considering a linear plant given by the Eq. (12) and the objective function is given in Eq. (13)

$$\frac{dx}{dt} = Ax + Bu \quad (12)$$

$$J(u) = \frac{1}{2} \int_0^T (x^T Q x + u^T R u) dt + \frac{1}{2} x^T(T) P_1 x(T) \quad (13)$$

where, Q , R , and P_1 are symmetric positive definite matrices.

$$H = x^T Q x + u^T R u + \lambda^T (Ax + Bu) \quad (14)$$

By maximizing the Hamiltonian Eq. (14), it gives an optimal solution in the form of $\lambda(t) = P(t)x(t)$. Then

$$\frac{d\lambda}{dt} = \frac{dP}{dt}x + P\frac{dx}{dt} = \frac{dP}{dt}x + P(Ax - BR^{-1}B^T P)x \quad (15)$$

$$-\frac{dP}{dt}x - PAx + PBR^{-1}BPx = Qx + A^T Px \quad (16)$$

This equation is satisfied by finding the $P(t)$ such that

$$-\frac{dP}{dt} = PA + A^T P - PBR^{-1}B^T P + Q \quad (17)$$

$$P(T) = P_1 \quad (18)$$

The matrix Q and R are obtained as

$$Q = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The value of controller gain K computed by $K = lqr(A, B, Q, R)$

$$K = \begin{bmatrix} 0.9650 & 0.1492 & 0.5617 & -1.5590 & 0.0703 & 0.2991 & 1.3661 \\ 0.1492 & 0.8912 & 0.0382 & 0.1120 & 0.2163 & 0.9542 & 4.5657 \end{bmatrix}$$

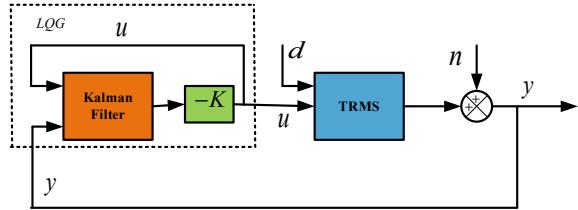
3.2 Linear PID Controller Design

The linear PID controller comprises three terms Proportional, Integral, and Derivative which represents the present errors, the accumulation of the past errors, and prediction of the future errors, respectively. The Eq. (19) represents the PID controller and system error.

$$u(t) = K_p[e(t) + \frac{1}{T_i} \int e(t)dt + T_d \frac{d}{dt}e(t)] \quad (19)$$

where $e(t) = r(t) - y(t)$, the Eq. (20) can be expressed in s-domain as Eq. (21).

$$U(s) = (P + \frac{I}{s} + D \cdot s) \cdot E(s) \quad (20)$$

Fig. 2 LQG controller

$$C(s) = \frac{U(s)}{E(s)} = \frac{Ds^2 + Ps + I}{s} \quad (21)$$

3.3 Linear Quadratic Gaussian (LQG) Controller Design

The linear quadratic Gaussian control is combined with LQR controller and linear quadratic estimator(Kalman Filter). A stochastic model is given in Eqs. (22) and (23) are considered to account for system nonlinearities such as process noise and sensor noise as viewed as white Gaussian noise, which disturbs the TRMS which is appeared in Fig. 2. The plant in this model encounters unsettling influences such as process noise n and sensor noise x which is driven by the control input u . The controller depends on the noisy estimations y to create these controls.

$$\dot{x} = Ax + Bu + Fd \quad (22)$$

$$y = Cx + n \quad (23)$$

where d the processes disturbance is vector and x is the sensor noise vector, which is assumed to be zero means Gaussian white noise with $E[d(t)] = E[n(t)] = 0$. The calculated disturbance and noise covariance are V_d and V_n are

$$v_d = \begin{bmatrix} -2.3299 \\ -1.4491 \\ 0.3335 \\ 0.3914 \\ 0.4517 \\ -0.1303 \\ 0.1837 \end{bmatrix} \quad v_n = \begin{bmatrix} -0.4762 \\ 0.8620 \end{bmatrix}$$

The state equation for the Kalman filter is given by Eq. (24). Where x_o the estimated state is vector and L is the Kalman filter gain. The state feedback control law is given in Eq. (25) and objective function is given in Eq. (26).

$$\dot{x}_o = (A - BK - LC)x_o + Ly \quad (24)$$

$$u = -Kx_o \quad (25)$$

$$J(u) = \int_0^{\infty} (x^T Q x + u^T R u) dt \quad (26)$$

The values of Kalman filter gain is computed by $L = \text{lqe}(A, B, C, 0.007 * B^T * B, C * C^T)$

$$L = \begin{bmatrix} 0.0001 & 0.0001 \\ -0.0000 & 0.0002 \\ 0.0000 & -0.0000 \\ 0.0003 & 0.0000 \\ 0.0000 & 0.0002 \\ 0.0000 & 0.0183 \\ -0.0000 & -0.0000 \end{bmatrix}$$

4 Results and Discussion

The experimental pitch and yaw responses of the TRMS with different reference step inputs by using LQR, L-PID, and LQG is shown in Fig. 3 and Fig. 4 respectively. The steady state and transient response of the TRMS in terms of settling time, overshoot and steady-state error of the TRMS are given in Table 2.

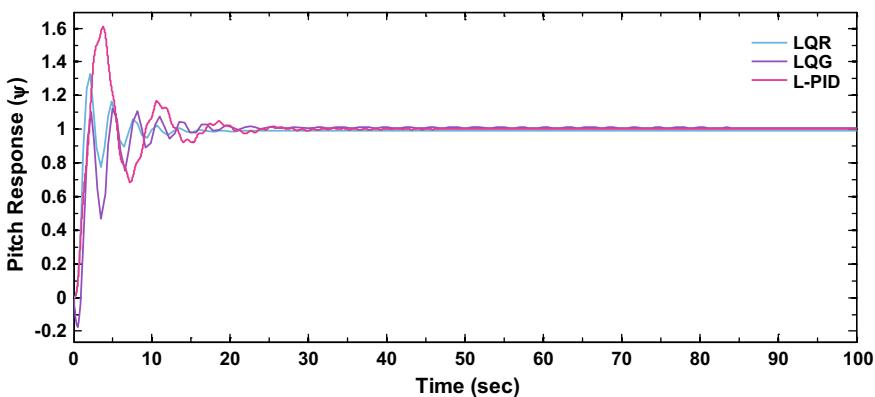
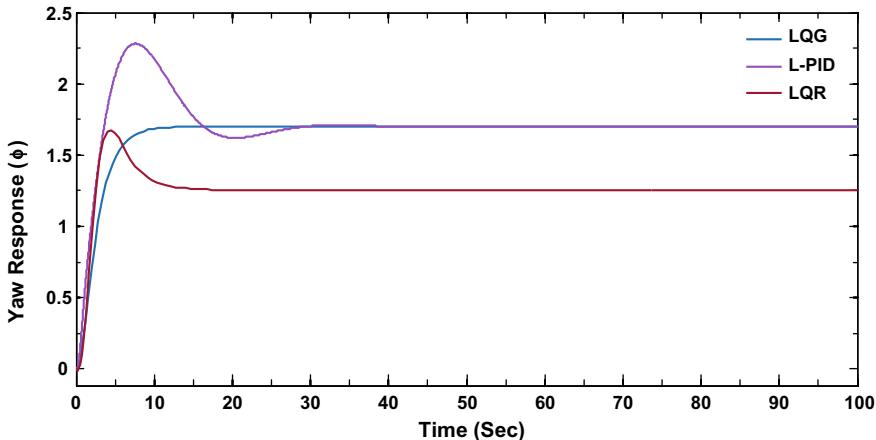


Fig. 3 Pitch responses

**Fig. 4** Yaw responses**Table 2** Performance parameters of TRMS using LQR, L-PID, and LQG controllers

Controller	Rotor	Reference value	Rise time (s)	Settling time (s)	Max overshoot (%)	Steady state error
L-PID controller	Yaw	1.7	2	45	70	0
	Pitch	1.0	3	28	40	0
LQR controller	Yaw	1.7	1	15	50	0
	Pitch	1.0	4	25	20	0
LQG controller	Yaw	1.7	6	8	0	0
	Pitch	1.0	3	15	15	0

5 Conclusion

In this paper, LQR, L-PID, and LQG controllers have been designed to control the lateral and longitudinal movement of TRMS and tested for their performance in terms of rise time, settling time, overshoot and steady state error. In the comparison, it is concluded that LQG controller has the advantage of minor overshoot, lesser settling time with the drawback of high rise time. From the outcomes, it is reasoned that the LQG gives better performance as far as both transient and steady state when contrasted with the LQR and L-PID controllers.

Acknowledgements The authors would like to thank the Govt. of India and Govt. NCT Delhi for providing funding through Delhi Technological University under TEQIP-II to procure TRMS Experimental Setup.

References

1. Feedback Instruments: Twin Rotor MIMO System (2016)
2. Sandino, L.A., et al.: A survey on methods for elaborated modeling of the mechanics of a small-size helicopter. Analysis and comparison. *J. Intell. Robot. Syst. Theory Appl.* **72**(2), 219–238 (2013)
3. Tastemirov, A., et al.: Complete dynamic model of the Twin Rotor MIMO System (TRMS) with experimental validation. *Control Eng. Pract.* **66**(8), 89–98 (2017)
4. Valluru, S.K., et al.: Comparative analysis of PID, NARMA L-2 and PSO tuned PID controllers for nonlinear dynamical system. *J. Autom. Syst. Eng.* **9**(2), 94–108 (2015)
5. Valluru, S.K., Singh, M., Mayank, S., Kumar, V.: Experimental validation of PID and LQR control techniques for stabilization of cart inverted pendulum system. In: 3rd IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology, pp. 708–712 (2018)
6. Wen, P., Li, Y.: Twin rotor system modeling, de-coupling and optimal control. In: IEEE International Conference on Mechatronics and Automation, pp. 1839–1842 (2011)

Improving Energy Efficiency of Hybrid ARQ Scheme for Cooperative Communication in UASNs Using PSO



Veerapu Goutham and V. P. Harigovindan

Abstract In this article, we present a mathematical model to compute the energy efficiency of hybrid ARQ scheme for cooperative communication (HARQ-CC) in underwater acoustic sensor networks (UASNs). Acoustic parameters such as frequency-dependent transmission losses, fading, and different noises present in the underwater channel are taken into account in this model. We propose a particle swarm optimization (PSO) technique to maximize the energy efficiency, by jointly optimizing the modulation level and packet size with respect to the distance between transceiving nodes. The analytical results show that the proposed PSO technique can significantly improve the energy efficiency of the HARQ-CC scheme for UASNs.

Keywords Cooperative communication · Energy efficiency · Particle swarm optimization (PSO)

1 Introduction

The applications of underwater acoustic sensor networks (UASNs) include ocean exploration, prediction of natural calamities, assisted navigation, naval warfare, underwater medical research, and many more [1]. However, UASN channel possesses unique characteristics, such as the time-varying multi-path propagation, Doppler spreading due to random movement of nodes, path loss dependent on the transmitted signal frequency, and ambient noises compared to terrestrial wireless sensor networks (WSN). Due to these underwater channel characteristics, the radio frequency and optical signals are inappropriate in UASNs. So, acoustic signals are commonly used in UASNs, which leads to limited bandwidth and longer propagation delays

V. Goutham (✉) · V. P. Harigovindan

Department of Electronics and Communication Engineering, National Institute of Technology
Puducherry, Karaikal 609609, India
e-mail: gouthamveerapu@gmail.com

V. P. Harigovindan
e-mail: hari@nitpy.ac.in

compared to WSNs. In UASNs, energy efficiency and reliability are the important objectives in designing a transmission protocol. The sensor nodes used in UASNs are powered by limited capacity batteries. It is difficult to recharge or replace the node batteries, if it gets drained due to the transceiving and sensing operations. Cooperative communication (CC) is a potential approach to limit the energy consumption of the sensor nodes by decreasing the number of retransmissions using the diversity links. In addition, UASNs comprise poor channel conditions which result in low reliability. The idea of combining CC with hybrid ARQ (HARQ) scheme improves the reliability and energy efficiency simultaneously. Here, the proposed scheme combines Reed–Solomon (R-S) codes with selective retransmission.

In this work, we present a mathematical model to calculate energy efficiency of HARQ scheme for cooperative communication (HARQ-CC) in UASNs. We evaluate the performance of energy efficiency by adopting the power consumption characteristics of WHOI practical underwater acoustic micro-modem. Further, we propose a particle swarm optimization (PSO) technique, which targets to maximize the energy efficiency by jointly optimizing the modulation level and packet size with respect to the distance between transceiving nodes. In [2, 3], authors proposed PSO-based clustering techniques and procedure for cluster head selection in WSNs and UASNs. In [4], authors proposed a scheme which uses PSO technique for solving localization problem in UASNs. PSO technique is extensively applied in cluster head selection, underwater node localization, deployment, and data fusion techniques [2–4], because it has a high potential to converge to the optimal solution. In this work, we apply PSO technique in two-variable optimization problem, which targets to maximize the energy efficiency. We consider packet size, modulation level as variables, and energy efficiency as the fitness function of the optimization technique.

2 Mathematical Model to Compute Energy Efficiency in UASNs

In this section, we present a mathematical model to evaluate the energy efficiency of direct, cooperative communication and HARQ-CC schemes in UASNs. We model underwater medium as Rayleigh fading channel due to the time-varying multi-path reception of signals at the receiver [5] and the gain provided by multiple paths is assumed to be independent and identically distributed (i.i.d) over packet transmission. We consider coherent M-ary phase shift keying (MPSK) as modulation technique. The closed-form approximation for finding the average symbol error rate (SER) of a Rayleigh fading channel between i th transmitting node and j th receiving node is given by [6]

$$SER_{ij} \approx \left[1 - \sqrt{\frac{\sin^2(\frac{\pi}{M}) SNR_{ij}}{1 + \sin^2(\frac{\pi}{M}) SNR_{ij}}} \right] \quad (1)$$

where SNR_{ij} is the signal-to-noise ratio of a link between the i th and j th transceiving nodes, respectively. The procedure to calculate the SNR_{ij} in UASN is presented in Sect. 2.1. The packet error rate (PER) of a link is given by

$$PER_{ij} = 1 - (1 - SER_{ij})^{\frac{X}{m}} \quad (2)$$

where X in bits is the packet size and $(1 - SER_{ij})^{\frac{X}{m}}$ gives the success probability of $\frac{X}{m}$ symbols transmitted to the receiver node. In the HARQ scheme, R-S codes pad the data packet with extra redundant check bits for error correction. So, the packet size is increased by an extra number of check bits (C). The average symbol error rate of a link after error correction is given by [7]

$$SER_{ij}^* \approx \frac{1}{2^K - 1} \sum_{k=t+1}^{2^K - 1} k \binom{2^K - 1}{r} SER_{ij}^k (1 - SER_{ij})^{2^K - 1 - k} \quad (3)$$

where the size of data packet is increased to $X + C = 2^K - 1$. The PER of a link after error correction (PER^*) is given by

$$PER_{ij}^* = 1 - (1 - SER_{ij}^*)^{\frac{X+C}{m}} \quad (4)$$

2.1 Calculation of SNR for Underwater Acoustic Link

The signal-to-noise ratio (SNR) of an underwater link is mainly governed by passive SONAR equation. It depends on transmitted source sound level (SL), transmission power loss (TL) comprises absorption and spreading losses, noise level (NL), and directivity index (DI). Here, DI is considered as 0 dB in the case of an omnidirectional antenna and 3 dB for the directional antenna case [5]. The overall SNR of an underwater link is given by

$$SNR_{dB} = (SL - TL - NL + DI)_{dB} \quad (5)$$

In shallow water scenario, the acoustic signals exhibit cylindrical spreading geometry due to the signals being bounded by the surface and bottom of ocean [5]. So, the transmitted source sound level related to the reference sound intensity is given by [5]

$$SL_{dB} = 10 \log \left[\frac{P_{tr}}{I_{ref} \times 2\pi r D} \right] \quad (6)$$

where P_{tr} is transmitted power of sound signal, $I_{ref} = \frac{q^2}{\rho \times c}$ is the reference sound intensity, D is the depth in meters, and r is the distance between transceiving nodes.

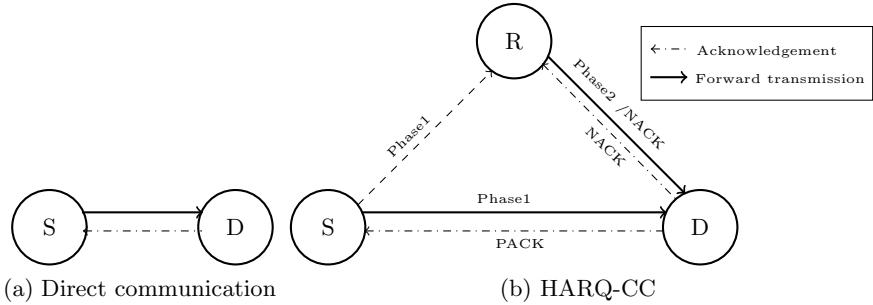


Fig. 1 Transmission schemes

TL offered by the underwater channel is given by [5]

$$TL_{dB} = k \times 10 \log r + r \times 10^{-3} \times 10 \log a(f) \quad (7)$$

where k is the spreading factor and $a(f)$ is the absorption coefficient given by Shulkin's and Marsh formula [8]. The value of k is 1 for cylindrical spreading geometry of signal propagation [5]. Noises present in underwater comprises turbulence, shipping, waves, and thermal noise. The overall NL provided by the underwater channel is given by [5]

$$NL_{dB} = 50 - 18 \log f \quad (8)$$

2.2 System Model

We consider a simple UASN model comprises source (S), destination (D), and relay (R) nodes as depicted in Fig. 1. The protocol description of general HARQ-CC is as follows: Initially, the node S broadcasts the data packets along with extra redundant bits generated by R-S codes for forward error correction. Due to broadcasting, the node R can also be able to receive the data packets along with the destination node. If the destination node D is not able to decode the data packet even after forward error correction, then it sends a negative acknowledgment (NACK) to the relay node R . Otherwise, it sends a positive acknowledgment (PACK) to the node S . If the relay node R decoded the data packet, then it transmits the data packet to the node D . Otherwise, it sends back a negative acknowledgment to the node D . Even after the retransmission by node R , if the node D is unable to decode the packet, then it will drop the data packet.

2.3 Energy Efficiency Model for HARQ-CC

Direct communication (DC) is a simple technique in which no relay nodes are available for cooperative communication as shown in Fig. 1a. Accordingly, the PER in DC technique is given by

$$PER_{DC} = 1 - (1 - SER_{sd})^{\frac{X}{m}} \quad (9)$$

The overall energy consumed in DC scheme is given by [9]

$$E_{DC} = (P_t + P_r) \frac{X}{R_b} \quad (10)$$

The first and second terms in Eq. (10) are the overall energy consumed for transmitting and receiving a data packet, respectively. The time taken to transmit a data packet is given by $\frac{X}{R_b}$ seconds. We define energy efficiency as the ratio of the total number of transmitted message bits to the overall energy consumption [9]. Mathematically,

$$\eta_{DC} = \frac{X_p(1 - PER_{DC})}{E_{DC}} \quad (11)$$

The detailed description of HARQ-CC is depicted in Fig. 1b. Accordingly, the PER in HARQ-CC is given by [9]

$$PER_{HARQ-CC} = PER_{sd}^* PER_{sr}^* + PER_{sd}^*(1 - PER_{sr}^*) PER_{rd}^* \quad (12)$$

Equation (12) shows that the unsuccessful packet transmission from source to destination nodes carried through either source to destination (S-D) or source to destination via relay node (S-R-D) paths. The overall energy consumed in HARQ-CC to transmit a data packet is described as follows [9]:

$$E_{HARQ-CC} = \left[\underbrace{(P_t + 2P_r)(1 - PER_{sd}^*)}_{\text{first term}} + \underbrace{(P_t + 2P_r)(PER_{sd}^* PER_{sr}^*)}_{\text{second term}} \right. \\ \left. + \underbrace{(2P_t + 3P_r)(PER_{sd}^*(1 - PER_{sr}^*))}_{\text{third term}} \right] \times \frac{X + C}{R_b} \quad (13)$$

The first and second terms in Eq. (13) are the energy consumed for successful packet transmission over S-D path and unsuccessful packet transmission over S-D and S-R paths, respectively, and the third term represents energy consumed for packet transmission over S-R-D path. The energy efficiency of HARQ-CC is given by

$$\eta_{HARQ-CC} = \frac{X_p(1 - PER_{HARQ-CC})}{E_{HARQ-CC}} \quad (14)$$

The overall improvement achieved in the energy efficiency of HARQ-CC scheme in comparison with the DC is computed using efficiency gain. It is given by [9]

$$G_{HARQ-CC} = \frac{\eta_{HARQ-CC}}{\eta_{DC}} \quad (15)$$

The overall PER (PER_{CC}), energy consumption (E_{CC}), and efficiency gain achieved in CC scheme (G_{CC}) are similar to HARQ-CC with reduction in the packet size to X bits.

3 Particle Swarm Optimization (PSO)

PSO technique is extensively applied in the several applications of WSNs and UASNs, namely, cluster head selection, node localization, deployment, and data fusion techniques. In this section, we present the PSO-based optimization technique, which targets to maximize the energy efficiency by optimizing the packet size and modulation. The energy efficiency depends on two vital parameters, packet size X , and modulation level m . In general, small size packets are less vulnerable to errors, while large size packets are more sensitive to packet errors, which reduce energy efficiency. A scheme with increasing modulation level carry more information per symbol but results in unnecessary packet errors. However, a low modulation level scheme is more robust to the packet errors, but it leads to an ineffective use of the channel. Accordingly, an optimization framework using PSO technique has been proposed by jointly varying the X and m with respect to the distance between transceiving nodes, which targets to maximize the energy efficiency. The step-by-step procedure of PSO-based HARQ-CC (PSO-HARQ-CC) is described as follows:

1. Initially, generate uniformly distributed swarm size of 1000 particles represented by $X_{(1,1)} \leq X_{(i,1)} \leq X_{(1000,1)}$ for the packet size and $X_{(1,2)} \leq X_{(i,2)} \leq X_{(1000,2)}$ for the modulation level, respectively.
2. Generate the random velocity vector, *i.e.*, $(V_{(1,1)}, V_{(1,2)}) \leq (V_{(i,1)}, V_{(i,2)}) \leq (V_{(1000,1)}, V_{(1000,2)})$ using uniform distribution in the range $(0, 1)$.
3. Compute the energy efficiency of HARQ-CC for all swarm particles and find global best (g_{best}) and personal best (p_{best}).
4. Update the velocity vector and swarm positions using

$$V_{(i,j)} = V_{(i,j)} + C_1 r_1 (p_{best_{(i,j)}} - X_{(i,j)}) + C_2 r_2 (g_{best_{(i,j)}} - X_{(i,j)}), \\ \text{for } 1 \leq i \leq 1000, 1 \leq j \leq 2 \quad (16)$$

$$X_{(i,j)} = X_{(i,j)} + V_{(i,j)}, \text{ for } 1 \leq i \leq 1000, 1 \leq j \leq 2 \quad (17)$$

Table 1 Parameters used for analysis

Parameter	Value
Transmitter power consumption (P_t)	48 W
Receiver power consumption (P_r)	3 W
Data rate	5 Kbps
Payload size X_p	41 bits
Packet size L	57 bits
Frequency f	25 KHz
Depth of shallow water	50 m
Modulation level	2 bits/symbol

where C_1 and C_2 are acceleration factors and r_1 and r_2 are uniformly distributed random numbers in the range $[0, 1]$.

5. Repeat the procedure from step 3 until the global optimum values are obtained.

4 Analytical Results

In this section, the analytical results for energy efficiency of DC, CC, HARQ-CC, and PSO-HARQ-CC schemes are presented and evaluated using MATLAB® R2016b. Table 1 presents the power consumption characteristics of WHOI micro-modem and other parameters used in this model [10]. In this work, the relay node is considered to be present exactly at the center of source and destination nodes in CC and HARQ-CC schemes. It has already proven that the energy efficiency of these schemes is optimum when the relay node is exactly at the center of the transceiving nodes in WSNs and UASNs [9, 11].

Here, we present a comparative analysis for energy efficiencies of DC, CC, HARQ-CC, and PSO-HARQ-CC schemes. Figure 2 shows that HARQ-CC scheme provides higher energy efficiency and efficiency gain than other schemes (DC, CC) for distance separation beyond 300 m between the transceiving nodes. It is also observed from Fig. 2 that the energy efficiency of HARQ-CC can be further increased using the proposed PSO optimization technique. From Fig. 2, it is clear that HARQ-CC with PSO technique (PSO-HARQ-CC) is having higher energy efficiency than HARQ-CC, CC, and DC techniques. It is found that PSO-HARQ-CC provides much higher efficiency gain compared to other schemes, when the distances between source and destination nodes are large.

Here, the HARQ-CC scheme is evaluated for fixed packet size ($X = 63$ bits) and modulation level ($m = 2$). For the PSO-HARQ-CC, packet size and modulation level are obtained using the PSO technique. The energy efficiency of PSO-HARQ-CC scheme has been optimized by jointly optimizing the packet size and modulation level. Accordingly, we presented the optimal values of modulation level and packet

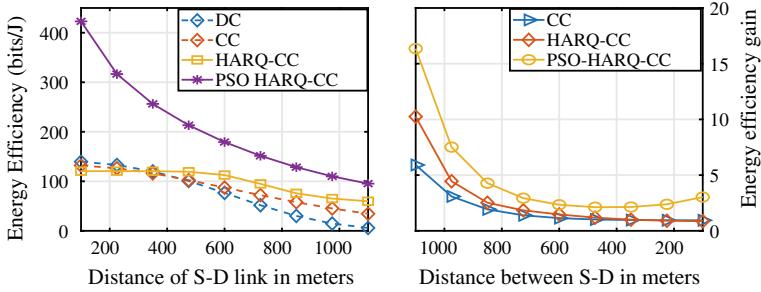


Fig. 2 Energy efficiency and gain versus the distance between S-D link

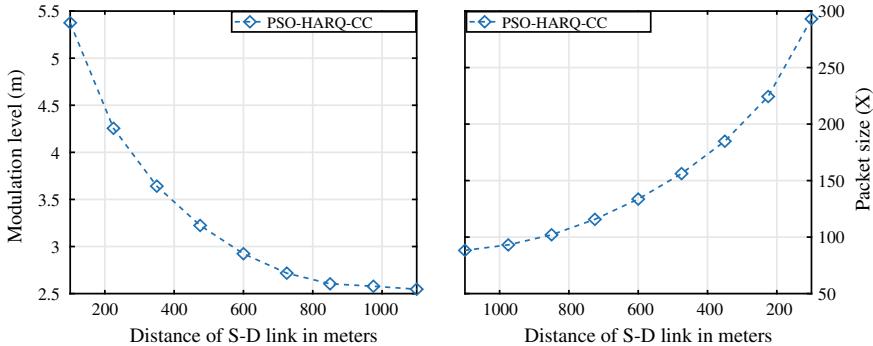


Fig. 3 Modulation level and packet size versus the distance between S-D link

size computed using the PSO technique in Fig. 3. The PSO-HARQ-CC scheme uses larger packet size and higher modulation level for the smaller distances between the transceiving nodes and further alters to smaller packet size and lower modulation level for the greater distances between the transceiving nodes to optimize energy efficiency.

5 Conclusion

In this article, a mathematical model to compute energy efficiency for different schemes (DC, CC, HARQ-CC) in UASNs is presented. We have performed extensive numerical analysis by considering the power consumption characteristics of WHOI practical underwater acoustic micro-modem. The analytical results show that HARQ-CC provides higher energy efficiency compared to the DC and CC schemes. We propose PSO technique to maximize the energy efficiency by jointly optimizing the packet size and modulation level. From the results and discussions, it is evident that HARQ-CC with PSO optimization algorithm (PSO-HARQ-CC) can

significantly improve the energy efficiency of UASNs. The presented work can be extended by integrating full duplex relay-aided communications, virtual MIMO (V-MIMO), and selecting optimal relay for cooperation as a future scope of the work.

References

1. Coutinho, R.W.L., Boukerche, A., Vieira, L.F.M., Loureiro, A.A.F.: Underwater wireless sensor networks: A new challenge for topology control-based systems. *ACM Comput. Surv.* **51**(1), 19:1–19:36 (2018)
2. Rao, P.C.S., Jana, P.K., Banka, H.: A particle swarm optimization based energy efficient cluster head selection algorithm for wireless sensor networks. *Wirel. Netw.* **23**(7), 2005–2020 (2017)
3. Azharuddin, M., Jana, P.K.: Pso-based approach for energy-efficient and energy-balanced routing and clustering in wireless sensor networks. *Soft Comput.* **21**(22), 6825–6839 (2017)
4. Zhang, Y., Liang, J., Jiang, S., Chen, W.: A localization method for underwater wireless sensor networks based on mobility prediction and particle swarm optimization algorithms. *Sensors* **16**(2) (2016)
5. Domingo, M.C., Prior, R.: Energy analysis of routing protocols for underwater wireless sensor networks. *Comput. Commun.* **31**(6), 1227–1238 (2008)
6. Goldsmith, A.: *Wireless Communications*. Cambridge University Press, New York (2005)
7. Sklar, B.: *Digital Communications: Fundamentals and Applications*. Prentice-Hall Inc, USA (1988)
8. Marsh, H.W.: Progress in underwater acoustics. *IEEE Trans. Sonics Ultrason.* **22**(5), 333–335 (1975)
9. Wang, S., Nie, J.: Energy efficiency optimization of cooperative communication in wireless sensor networks. *EURASIP J. Wirel. Commun. Netw.* **2010**(1), 162326 (2010)
10. Yildiz, H.U., Gungor, V.C., Tavli, B.: Packet size optimization for lifetime maximization in underwater acoustic sensor networks. *IEEE Trans. Ind. Inform.* **1** (2018)
11. Geethu, K.S., Babu, A.V.: Minimizing the total energy consumption in Multi-hop UWASN. *Wirel. Pers. Commun.* **83**(4), 2693–2709 (2015)

Implementation of FM-Based Communication System with 3-Level Parallel Multiplier Structure for Fast Transmission Using FPGA



Salauddin Mohammad, R. Seetha, S. Jayamangala, B. Khaleelu Rehman and D. Kumar

Abstract Any communication setup with the two individual hardware gives the best measure of real-time parameters like hardware, speed, total power consumption, etc. The research article deals with the implementation of a 3-level parallel multiplier structure, which can be used mainly for FM Broadcasting system. The proposed system has been implemented using communication between the two FPGAs. The message signal with 1 MHz frequency, 16-bit sample signal is generated by the Direct Digital Synthesizer (DDS) and is modulated with 10 MHz, 16-bit sampled carrier signal and transmitted over the channel. At the receiver, the signal is passed through a noise removal FIR filter and demodulated for the reconstruction of the original signal. The target of the design has included area and speed optimization as well. The system has been simulated on Xilinx ISE 14.7 and ModelSim 10.1 software coded in VHDL. Synthesis reports are generated from ChipScope Pro Tool with ARTIX-7 board FM communication analysis. The simulated resulted are verified for the FM range 88–108 MHz.

Keywords Software-Defined Radio (SDR) · Field-Programmable Gate Array (FPGA) · Frequency Modulation (FM) · Direct digital synthesizer (DDS)

S. Mohammad · R. Seetha (✉) · S. Jayamangala
ECE Department, J.B. Institute of Engineering and Technology, Hyderabad, India
e-mail: rseetha.ece@jbiet.edu.in

S. Mohammad
e-mail: salal.vlsi@gmail.com

S. Jayamangala
e-mail: jayamangala.ece@jbiet.edu.in

B. K. Rehman · D. Kumar
University of Petroleum & Energy Studies, Dehradun, India
e-mail: krehman@ddn.upes.ac.in

D. Kumar
e-mail: d.kumar@ddn.upes.ac.in

1 Introduction

Software-Defined radio (SDR) momentous inventions for extreme prototyping of FM radio system [1] using reconfigurable hardware platforms, which provide upper hand on conventional analog and hardware-oriented method. There are many SDR prototyping platforms, which have been discussed by USRP and Microsoft Sora [2]. The selection of the SDR platform for implementing and developing the components of complex SDR system can be done by referring many of the proposed architectures in [3–5]. DDS has been chosen instead of Voltage-Controlled Oscillator (VCO) since the digital implementation of the FM gives efficient results in terms of speed power consumption and also broadcasting standards. In the presented paper, multiplier structure also addresses the speed improvement over the normal multiplication. 16-bit multiplication has been chosen for the verification on ARTIX-7 FPGA. The DDS has been used to generate the baseband and the carrier signals of the FM onto which the multiplier structure has been applied [6, 7].

The paper is organized as follows: Sect. 2 describes the FM modulator and demodulator principles and architecture. The proposed 3-level multiplier structure and operation are discussed in Sect. 3 followed by the results and discussions in Sect. 4. Section 5 concludes the paper.

2 FM Modulation and Demodulation

In the transmitter section, modulation process takes place where the frequency of the carrier signal is varied in accordance with the message signal. This can be achieved by multiplying the carrier by the message signal [8]. Two 16-bit signals with frequencies 1 and 10 MHz are chosen as the message and carrier signal. Later on, the frequency has been upconverted by using an upconverter. DDS generates the sine wave from the lookup table data. The signal equation and the block diagram of the digital FM transmitter are given in Fig. 1.

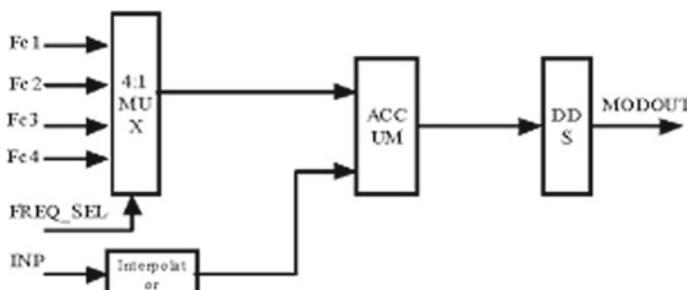


Fig. 1 Block diagram of Digital FM Transmitter

$$S_{FM}(t) = A_c \cos 2\pi [f_c t + k_f \int_0^t m(p) dp]$$

where A_c and F_c represent the amplitude and frequency of the carrier and K_f being the frequency deviation constant.

In the transmitter, the mixer work is done by the multiplier. A multiplier is a device which has two input signals and gives one output signal which is proportional to the product of amplitude of two input signals.

$$V_{out} = K \cdot Vin_1 \cdot Vin_2$$

where K is a constant with the dimension of $1/V$.

Multipliers are used for analog computation, such as speed and power in a circuit. The architecture of multiplier can be divided into two possibilities: (i) Parallel multiplier and (ii) constant coefficient. Parallel multiplier takes two inputs of the same bit length and gives the product of these two values which results in a tradeoff between dedicated multiplier resources, slice logic and maximum achievable clock frequency [9]. Constant coefficient multiplier takes one input and it is multiplied to a user-defined constant value. The multiplier can be constructed from block memories, distributed memory in conjunction with slice logic [10, 11].

3 Proposed 3-Level Multiplier Architecture

In this paper, the ordinary multiplier has been replaced by the mixer with a 3-level multiplier to optimize the area and speed of transmission. This uses dedicated multiplier primitives and lookup tables for implementation of the multiplier. IP core will use Xtreme DSP Slice and MULT 18 * 18 for area optimization. It will use both LUT4- and LUT6-based FPGA devices for speed optimization. This operation is done by pipelining and control signals. For maximum speed optimization, 3-level pipelining has been chosen. These three levels are implemented by treating as registers. Pipeline level 0 is used for combinational logic circuitry implementation. Pipeline level 1 is used for output registers implementation. Pipeline level 2 is used to insert between input and output to improve achievable clock speed while increasing latency (Fig. 2).

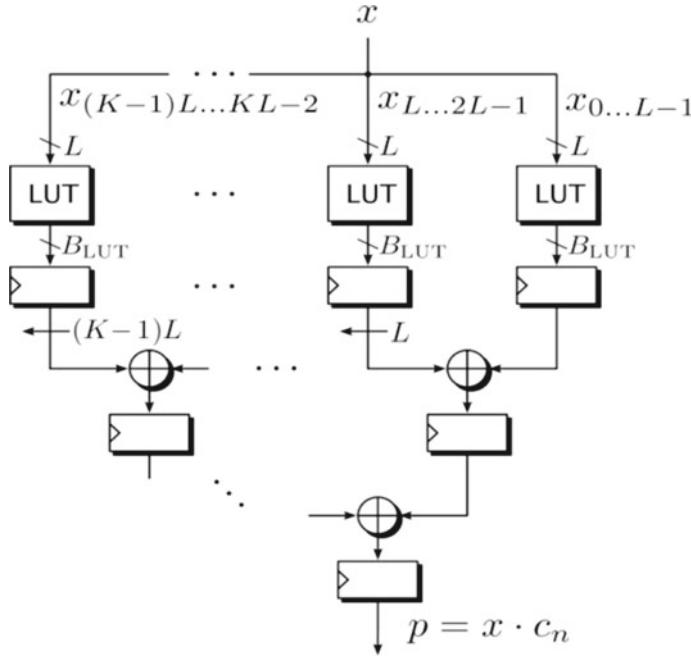


Fig. 2 Lookup table based 3-level parallel multiplier architecture

4 Hardware Implementation Details and Discussions

A. Synthesis Results:

Artix-7 FPGA device has been used as the target device for Implementation and Xilinx 14.7 has been used as a synthesis tool for FM Transmitter and the timing parameters with and without optimization are listed below. The result concludes that Maximum output required time after the clock is reduced greatly with 3-level multiplier architecture as shown in Table 1.

B. RTL diagram:

The RTL schematic and its internal architecture are shown in Figs. 3 and 4.

C. ChipScope tool results:

For the post-place and route simulation in FPGA 14.7—Xilinx ISE is used as a simulator. A bit stream of the sine wave of 1 MHz with a 29-bit phase increment is given to the FM transmitter to transmit the signal. At the receiver, the signal is decoded and multiplied with other 1 MHz signal. The required frequency from the DDS is generated by using the following formula:

$$pinc_in = \frac{Freq}{Fclk} * 2^{pinc_in\ bits}$$

Figures 5, 6 and 7 represent the outputs of modulation and demodulation of FM modulation with and without optimization in which cosine signals of 1 and 10 MHz are represented.

5 Conclusion

The research article presents the implementation of FM Broadcasting system using two FPGA's of Nexys 4 DDR Artix 7 boards. A 1 MHz, 16-bit sample baseband signal and 10 MHz, 16-bit sampled carrier signal was generated by the DDS and modulated with 3-level pipeline multiplier IP core and transmitted over the channel. At the receiver, the signal is passed through a noise removal FIR filter and demodulated for the reconstruction of the original signal. The target of the design area and speed optimization was achieved as shown in the results. The system has been simulated

Table 1 Timing reports

S. No	Parameters	Without optimization	3-level multiplier
1	Minimum period: (Maximum Frequency: 209.927 MHz)	4.764 ns	4.660 ns
2	Minimum input arrival time before the clock	4.060 ns	4.060 ns
3	Maximum output required time after the clock	4.749 ns	0.398 ns
4	Maximum combinational path delay	No path found	No path found

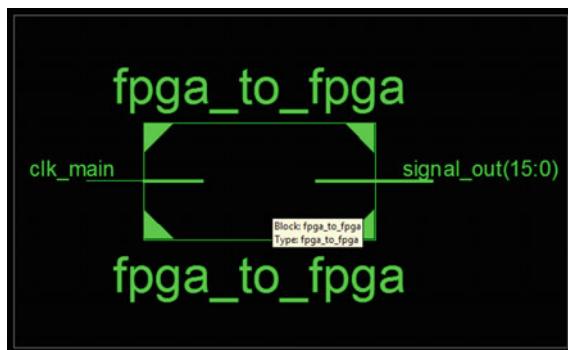


Fig. 3 RTL schematic

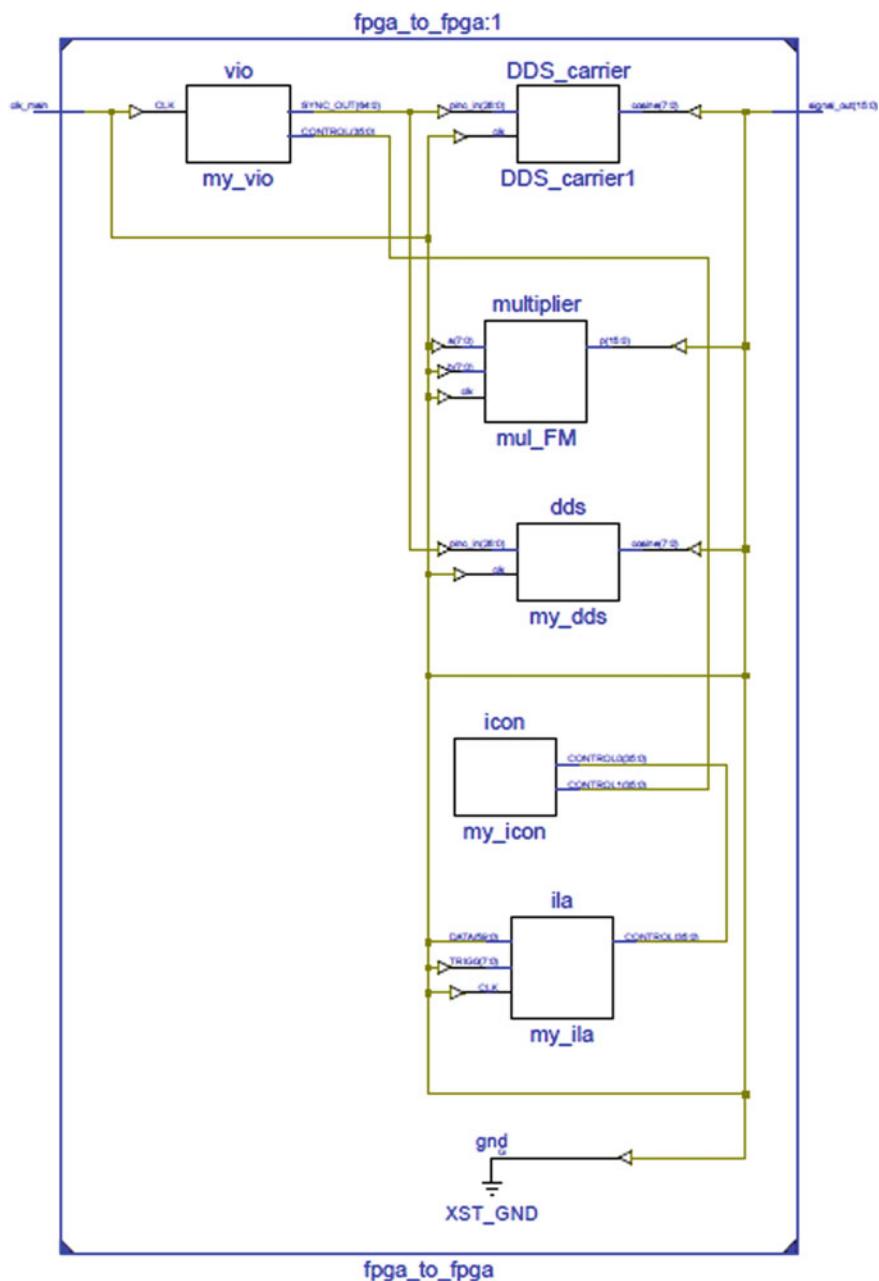


Fig. 4 RTL schematic internal structure

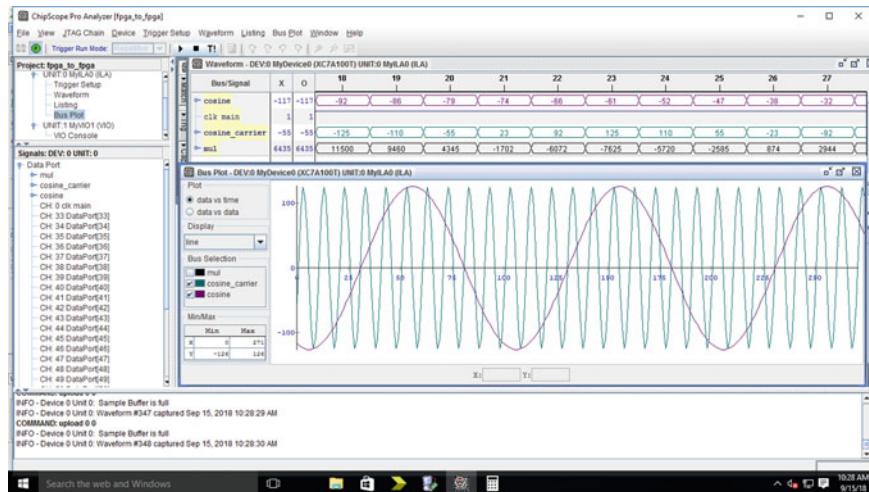


Fig. 5 Cosine wave of message and carrier without optimization

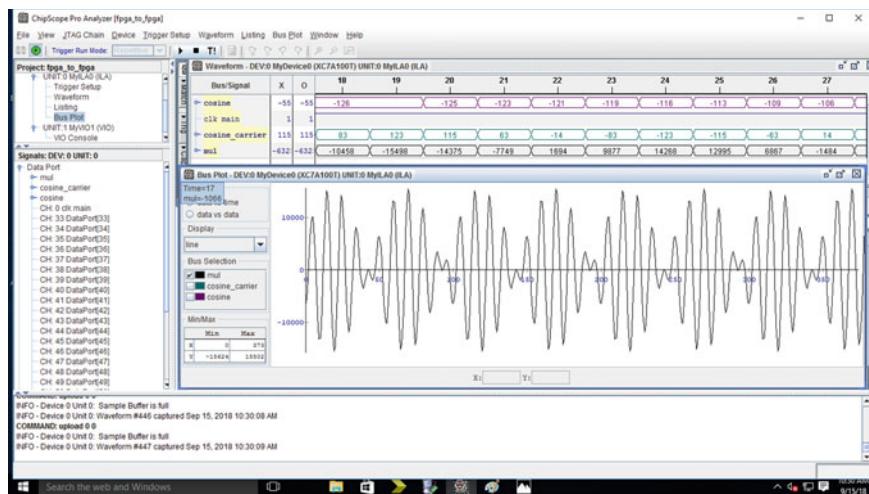


Fig. 6 Modulated waveform with 3-level multiplier

on Xilinx ISE-14.7 and ModelSim 10.1 software with the help of VHDL language. ChipScope Pro tool is used for the FM transmitter analysis. The simulated resulted are verified for the FM range 88–108 MHz.

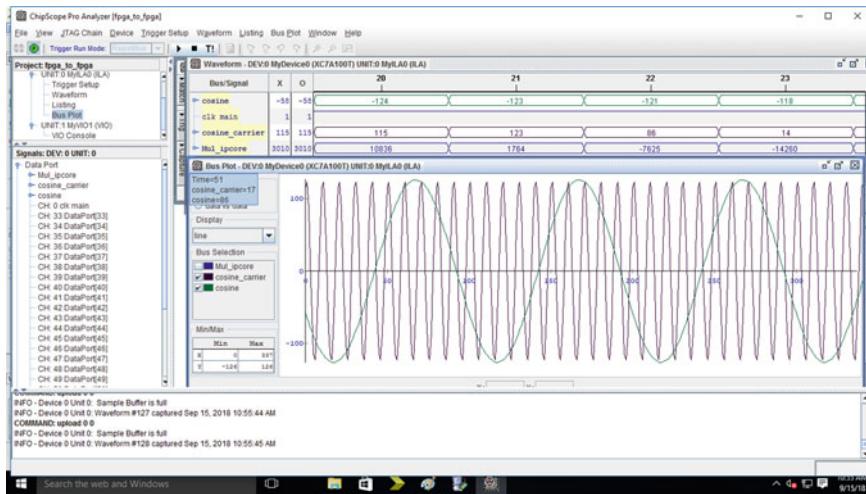


Fig. 7 Demodulated cosine of message and carrier

References

1. Hatai, I., Chakrabarti, I.: FPGA implementation of a digital FM modem. In: 2009 International Conference on Information and Multimedia Technology, Jeju Island, pp. 475–479 (2009)
2. Zhang, C., Anderson, C.R., Athanas, P.M.: All digital FPGA based FM radio receiver. SDR Forum Technical Conference (2007)
3. Fubing, Y.: FPGA implementation of a fully digital FM demodulator. In: The Ninth International Conference on Communications Systems, 2004. ICCS 2004, Singapore, China, pp. 446–450 (2004)
4. Brito, J.P.M., Bampi, S.: Design of a digital FM demodulator based on a 2nd-order all-digital phase-locked loop. *J. Analog Integr. Circ. Signal Process.*, Springer Netherland, May 28 2008, Numbers 1–2/November 2008
5. Song, B.S., Lee, I.S.: A digital FM demodulator for FM, TV, and wireless. *IEEE Trans. Circ. Syst. II Anal. Digit. Signal Process.* **42**(12):821–825
6. Indranil, H., Indrajit, C.: A New High-Performance Digital FM Modulator and Demodulator for Software-Defined Radio and Its FPGA Implementation. Received 31 March 2011; Accepted 8 September 2011
7. Datasheet of CORDIC, DAC, ADC, DDS
8. Brito, J.P.M., Bampi, S.: Design of a digital FM demodulator based on a 2nd order all-digital phase-locked loop. In: Proceedings of the 20th Annual Conference on Integrated Circuits and Systems Design, pp. 137–141 (2007)
9. Brito, J.P.M., Bampi, S.: Design of a digital FM demodulator based on a 2nd-order all-digital phase-locked loop. *J. Anal. Integr. Circ. Signal Process.* Springer, Netherland, no. 1–2/November 2008, May 2008
10. Zhang, C., Anderson, C.R., Athanas, P.M.: All digital FPGA based FM radio receiver. In: SDR Forum Technical Conference (2007)
11. Yu, F.: FPGA Implementation of a fully digital FM demodulator. In: Proceeding of the Ninth International Conference on Communications Systems ICCS, pp. 446–450 (2004)

Virtual Simulation of Regulated Power Supply for Various Microcontroller Boards and Their Peripherals



Varnita Verma, Piyush Goyal, C. S. Meera, Mukul Kumar Gupta and Piyush Chauhan

Abstract The DC regulated power supply is required to power various microcontroller boards and different sensory devices. Microcontrollers are widely getting used in upcoming new electronic gadgets. This paper describes the circuit diagram, PCB design, and footprints of the components used. The circuit simulation provides the output with respect to input in circuit, which helps in validation of circuit. The 3D virtual simulation helps the researcher to analyze the component placement along with the 3D view of PCB in simulation environment. The TINA software used as simulation software to analyze the circuit.

Keywords Regulated power supply · PCB design · Virtual simulation

1 Introduction

A microcontroller is a system containing a processor and embedded system. Microcontrollers are widely getting used in upcoming new electronic gadgets. The basic necessary power supply of a microcontroller is 12 V used in Raspberry Pi, Arduino,

V. Verma (✉) · P. Goyal · C. S. Meera · M. K. Gupta

Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India
e-mail: varnitaverma2@gmail.com

P. Goyal
e-mail: piyushgoyal1408@gmail.com

C. S. Meera
e-mail: meerasunil007@gmail.com

M. K. Gupta
e-mail: mkgupta@ddn.upes.ac.in

P. Chauhan
Department of Analytics, School of Computer Science, University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India
e-mail: pchauhan@ddn.upes.ac.in

and various microcontrollers. Though 220 V is quietly a high voltage for a microcontroller, it must be converted to 12 V because fluctuation in AC voltage may burst the microcontroller [1]. In order to extract 12 V DC power supply from 220 V AC supply, a 12 V regulated power supply is made for different controller circuits so as the fluctuation of AC voltage does not lead to any destruction of any microcontroller. 12 V DC regulated power supplies are one of the most common power supplies in use today. A 12 V power supply can be obtained from either 220 or 240 V power supply. It is generated with help of diodes, transformers, etc.

Power supplies can be of two types:

1. Regulated Power Supply,
2. Unregulated Power Supply.

Unregulated Power Supply as the name indicates, they do not provide clean voltage. Here, clean voltage means constant voltage. Moreover, Regulated Power Supply has voltage regulate at the output in order to extract stabilized voltage.

The typical applications for 12 V DC power supplies are the following:

- Computer-based device and used in various networks,
- Telecom communications and Fiber optic Network,
- Analog Communications,
- Instrumentation Engineering,
- Power Industries,
- Medical industries,
- Motor control,
- Structural monitoring.

So, a 12 V power supply is widely used for many purposes [2, 3].

Growing demand of 12 V power supply leads to construct such a type of circuit, which will enhance all the abovementioned applications.

1.1 Requirement of Power Supply

There are various issues, which may happen with the approaching force supply to a site. These issues cause voltage and current fluctuation, which can significantly affect gear activity and power utilization [4, 5]. Voltage security alludes to the capacity of a power framework to keep up relentless voltage levels after an unsettling influence. The primary types of voltage insecurity can be defined as:

1. Managed low or High Voltage,
2. Waveform twisting.

1.1.1 High/Low Voltage

High and low voltage may lead to create a disturbance in power supply. Less voltage in a power supply may lead to the losses in equipment [6, 7]. Moreover, excessive voltage may lead to burning of equipment in the circuit.

1.1.2 Waveform Distortion

Waveform distortion can cause data issues and equipment's related problems.

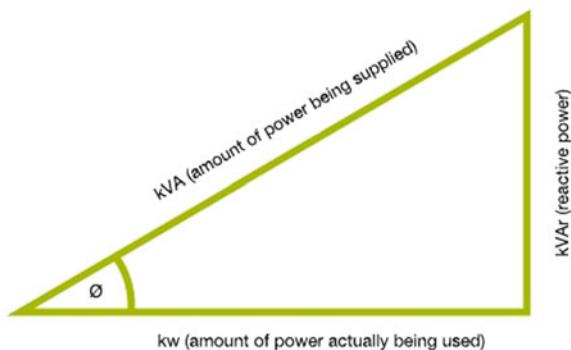
The circuit is designed in such a way that it converts high-voltage AC power supply to low-voltage DC power supply.

1.2 Effect on Hardware

Insufficient power supply on a hardware can lead to shorten the life of the hardware resulting in unstable operations causing failure [6, 8].

- **Motors:** When motors are given insufficient power supply, they may not work well sufficiently and also cause malfunction. Under voltage may cause overheating resulting in shortening the life of motor.
- **Sensitive equipment's:** When disturbed power supply is provided to sensitive electronic gadgets, the gadgets may not work properly.
- **Impact on bills:** Customers pay the bills of the electricity of what they had not used actually. The power supplied to them is high. The power used by them is low. In order to retain less charge on electricity bills, adequate supply is necessary. The power triangle shown in Fig. 1 indicates the variations between the actual power sent to the customer and the power consumed by the customer.

Fig. 1 Power triangle



1.3 Methods to Improve Power Quality

- For rectification of power quality, power factor correction (PFC) should be taken into practice.
- Optimization of voltage can help in stabilizing power and leads to reduce the impact of voltage fluctuation.

2 Methods and Material

Power Issues on electrical system causes interruption in hardware interfacing by providing operational delay, gear precariousness and disappointment in results. Poor power quality additionally results in higher losses as expected because of improper utilization of power framework. A poor power supply can cause the system breakdown. In order to avoid such condition circuit flow chart and different stages of signal regulation were discussed below.

The power circuit design flowchart can be seen in Fig. 2, which consists of four stages through which our 220 V and 50 Hz mains is converted into required 12 V regulated DC supply. These four stages are defined below, and each stage fulfills a specific purpose.

1. AC-to-AC Conversion,
2. AC-to-DC Conversion—Full-Wave Rectification,
3. Smoothing,
4. Voltage Regulation.

2.1 Circuit Design

The circuit is designed using in such a way that each component in the circuit is responsible for any function. The circuit simulation design can be seen in Fig. 3.

In order to make 12 V DC power supply from 220 V AC power supply, high voltage (i.e., 220 V) needs to be step down. A step down transformer is used with an



Fig. 2 Flowchart of regulated DC supply

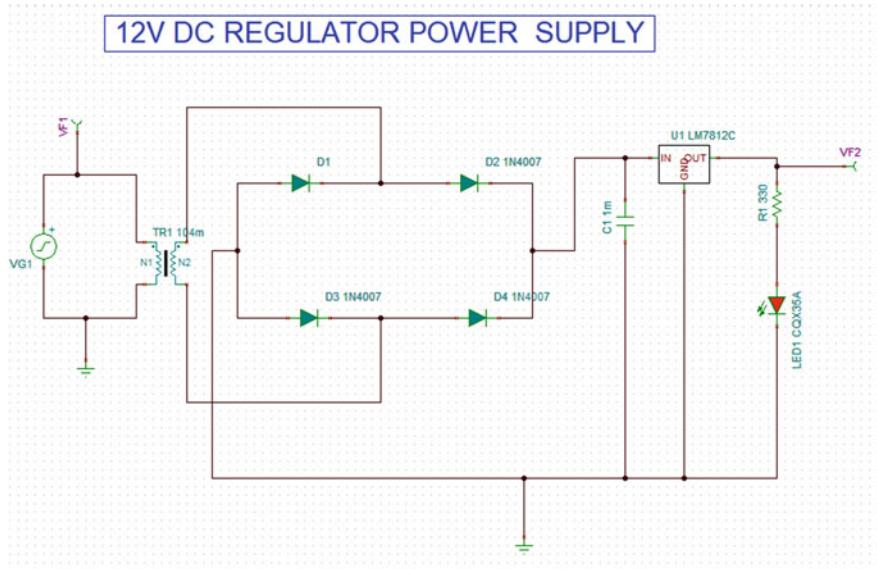


Fig. 3 Circuit simulation in TINA software

input supply of 220 Vs and its secondary coil is connected to the developed circuit. The circuit consists of full-wave bridge rectifier, capacitor, voltage regulator, and load resistance with LED as an indicator device. The rectifier is connected to four 1N4007 diodes, namely D1, D2, D3, and D4. The negative end of diode D1 and positive end of diode D2 are connected to one end of the secondary coil. The negative end of diode D4 and positive end of diode D3 are connected to other end of the secondary coil. The negative ends of D1 and D3 are connected to that terminal from which output of rectifier can be drawn out. Positive ends of diodes D2 and D4 are connected to that terminal from which output of full-wave rectifier can be drawn out. A capacitor of 1000μ is connected between the output terminals to create a ripple-free DC voltage. LM-7812 IC is connected in parallel to 1000μ capacitor to regulate voltage. The output voltage can be taken out from voltage output terminal of LM-7812 IC. The LED glows will indicate that the power supply is on and working properly.

The schematic diagram for PCB design is shown in Fig. 4, which clearly provides pictorial representation of components placement and connection on PCB. This helps in getting information of size of PCB, depth of drill hole, Number of holes, etc., as shown in Fig. 5.

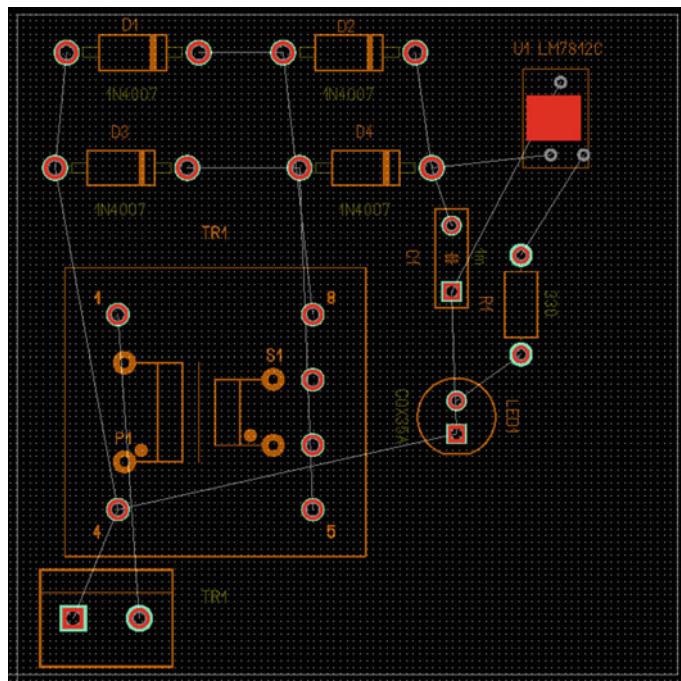


Fig. 4 Schematic diagram of PCB

```
PCB Information - X

Project name: C:\Users\Administrator\Desktop\UU1 - AUTOSAVE 18-10-24 16_14.tpc
Schematics name: C:\Users\Administrator\AppData\Local\Temp\DesignSoft\{27080F8C-2F94-4E6A-A41F-93832EB95132}\UU1 - AUTOSAVE 18-10-24 16_14.TSC

Number of components:
    Total    Thru-hole    SMD
Total      11          11        0
Top        11          11        0
Bottom     0            0         0

Number of nets: 12
Number of connections: 16 (0 routed, 16 unrouted)
|
Number of track segments: 0
Number of vias: 0

Number of drill holes: 27
Smallest diameter: 20.00 mil
Largest diameter: 44.00 mil
```

Fig. 5 PCB fabrication parameters

3 Results and Discussions

The regulated power supply is achieved on implementing the above circuit on PCB. The input signal AC signal of 220 V is converted to 12 V DC regulated power supply. In Fig. 6, VF1 shows the AC input waveform and VF2 as a 12 V DC regulated supply, which is very useful to power up the microcontroller boards and their peripheral devices.

In Fig. 7, three-dimensional virtual visualization view enables the designer to intuitively get the framework architecture and can further be used to reduce the size of PCB.

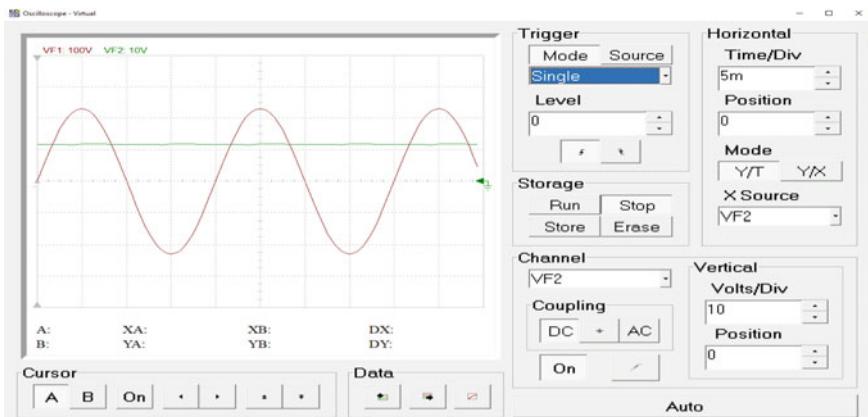


Fig. 6 Input and output voltages of regulated power supply circuit



Fig. 7 3D virtual visualization of PCB

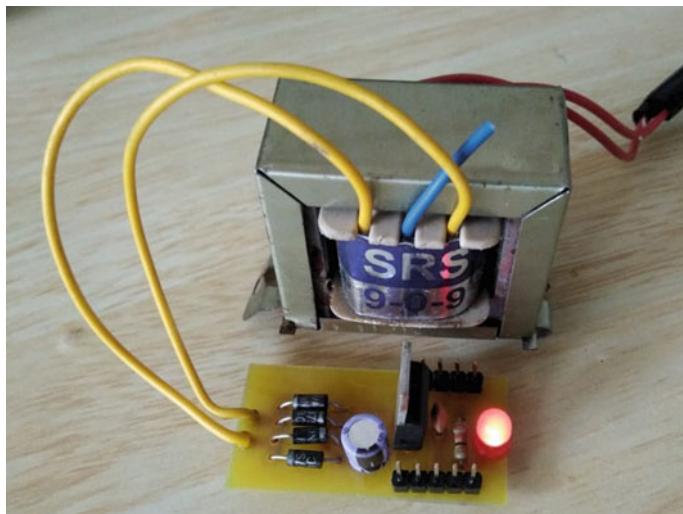


Fig. 8 Hardware construction of 12 V regulated power supply

Hence by doing all the circuit connections properly, we can have a 12 V DC supply. A constant power supply is obtained from the hardware constructed as shown in Fig. 8 and given to microcontrollers for their respective applications.

References

1. Nejabatkahah, F., Li, Y.W.: Overview of power management strategies of hybrid AC/DC microgrid. *IEEE Trans. Power Electron.* **30**(12), 7072–7089 (2015)
2. Verma, V., Kumar, R., Kaundal, V.: Implementation of ladder logic for control of pipeline inspection robot using PLC. In: Singh, R., Choudhury, S. (eds.) Proceeding of International Conference on Intelligent Communication, Control and Devices. Advances in Intelligent Systems and Computing, vol 479. Springer, Singapore
3. Verma, V., Mishra, R.: Applications of wireless sensor nodes in structural health monitoring: a review. *HCTL Open Int. J. Technol. Innov. Res. (IJTIR)* **18**(2). <http://ijtir.hctl.org> (2016). (e-ISSN: 2321-1814, ISBN (Print): 978-1-944170-16-5)
4. Zakaryukin, V., Kryukov, A., Cherepanov, A.: Intelligent traction power supply system. In: Energy Management of Municipal Transportation Facilities and Transport, pp. 91–99. Springer, Cham (2017)
5. Llona, G.F.: System for determining the nominal voltage of a power supply. US Patent 9,585,194, issued Feb 28, 2017
6. Chuang, T.-J., Wu, S.-H.: Power supply system and voltage output module. US Patent Application 15/867,711, filed July 26, 2018

7. Owen, D.K., Gill, J.S., Okamura, K., Toma, S., Tallak, S.: Performing a health check on power supply modules that operate in a current sharing mode. US Patent Application 10/082,856, filed Sept 25, 2018
8. Yamada, K., Matsui, R., Yagi, A., Yoshimi, N.: Power supply system. US Patent 8,310,094, issued Nov 13, 2012

Contribution of Learner Characteristics in the Development of Adaptive Learner Model



Amit Kumar and Vishal Bharti

Abstract Learner models are built to enable adaptive and personalized instructions to prospective learners in Adaptive Intelligent Tutoring Systems (AITS). They are created parallel to the advancement of adaptive intelligent tutoring frameworks and consequently, connected to the system's advancement. This paper presents the critical literature review from 2002 to 2017 in the research area of learner characteristics in the development of a learner model for the particular field of AITS. The primary aim of this literature review is to answer the two fundamental questions. First, what are the main characteristics of learners to be modeled second, what is the contributions of learner characteristics in the development of the learner model? The learner characteristics model along with a model application and their comparative study has been discussed in ITS. The goal of this paper is to give the gained information to educationalists, developers, researchers, and scholars for making decisions about what the learner's characteristics are should be followed while developing an effective learner model.

Keywords Learner model · Learner characteristics · Adaptivity · E-learning · Intelligent tutoring system

1 Introduction

Technological improvements cause constant changes to each part of the cutting edge society. The education itself couldn't stay inactive and unconcerned, all customary

A. Kumar (✉)

School of Computer Science and Engineering, University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India
e-mail: amitanit007@gmail.com

V. Bharti

Department of Computer Science and Engineering, DIT University, Dehradun 248009, Uttarakhand, India
e-mail: hod.cse@dituniversity.edu.in

encouraging tutoring systems are overhauled and reexamined and the new ones are presented [1, 2]. Toward this objective, different e-learning frameworks have been created amid the most recent years, be that as it may, a large portion of them shape static outdated applications, missing functionalities like instructive mixed-media conditions, customized abilities, and following of students' info and significance criticism. Customized bolster for students turns out to be much more critical when e-learning happens inside open and dynamic learning and data systems. In any case, the larger part of e-learning frameworks show the student as an element joined by a static predefined set of premiums and alternatives, without giving a suitable consideration regarding their requirements. Also, customized picking up utilizing appropriated data in powerful learning conditions is as yet an unsolved issue in e-learning research. A few methodologies toward versatile online instructive frameworks are at present examined, attempting to offer customized access and introduction offices to learning assets for particular application spaces. The absence of programmed assessment of student profile and the giving of fitting learning assets is a typical wonder in these frameworks. Furthermore, none of these frameworks make a factual investigation over the profiling in a request to improve their execution.

The learner model is usually named as user model or student model, learner models commonly store learner information (e.g., knowledge, individual preferences, learning strategy, and objective) and empower a framework to adjust its conduct to the individual learner [3, 4]. Learner models are utilized as a part of an assortment of frameworks, such as Adaptive Educational Systems, Pedagogical Recommender System, Mobile or Ubiquitous Systems, Education Hypermedia, Learning Management System, and ITS [5]. The learner model is a description of activities performed during learning and comprises of information about the learner or a vigil on learner action throughout learning. Ordinarily, a learner model contains the data about a learner such as his or her knowledge level, learning preferences, learning path, objectives, background, and interests [6].

Therefore, the essential motive is, development of an emphatic learner model using learner characteristics. The learner model alludes to the dynamic portrayal of the developing of knowledge, aptitude, and abilities of the learner. The learner model empowers the learning framework to adjust to the student who utilizes it and preferably incorporates all data about the student's conduct and information that highly influence their learning and execution [7]. The learner model content relies upon the learning condition and incorporates derived data about perspectives. Taking an example, a student's objectives, their plans, information, learning abilities, and attitudes provide the data that is collected. However, the most vital data about a student is his or her insight into the subject that is being examined [8].

No tutoring system can be replaced by a tutor system without the knowledge of the learner model. The adaptability and personalization feature is heavily dependent on the learner model, which is an essential part of ITS [1]. Most of specialists/pedagogues concur that an intelligent tutor ought to have a learner model [9–13]. There are several techniques to develop the learner model such as Overlay approach, the Stereotype modeling, the Perturbation model, Machine learning technique, Fuzzy learner modeling, Constraint-based modeling, Bayesian networks modeling, Cognitive theories, and Ontology-based modeling techniques. The detailed discussions of

modeling techniques are not in the scope of work. Therefore, the current research challenge is the development of an adaptive learner model in order to provide the adaptability and personalization in ITS. Learner modeling is the key factor that affects instructional decision [14].

2 Methods

This literature review is based on the learner characteristics used in the development of the learner model used in a specific field of Intelligent Tutoring System (ITS). In this paper, we discussed all the relevant and quality papers based on the comprehensive search that was published from 2002 to 2017. This research was terminated in the month of December 2017. The main criteria of inclusion in the literature review were that we included high impact journals and the search was limited to electronic databases such as Scopus, Science Direct, Springer links, Thesis, InderScience, and Dissertation. Scopus is one of the most accepted electronic search databases for all types of research papers. About 80% of journals based on the e-learning system/adaptive tutoring system/adaptive hypermedia were considered and listed in the Scopus search database. The journals that have low impact and low quality were excluded from this review. This literature review focused on the keywords namely “learner characteristics”, “learner model”, “learner modeling techniques”, “adaptive learner model”, “intelligent tutoring system”, “adaptive hypermedia”, “student modeling”, “E-learning System”, “Personalized Learning System”, “user modeling”, “student/user model”, “user modeling techniques in ITS” and was limited to only these keywords.

3 Literature Review

This literature review expects to contribute to determining the current research trends, circumstances, and gaps in the literature. This critical analysis helps the researchers and educational system developers for selecting the relevant learner modeling techniques in the field. This study presents an idea about the current parameters/characteristics of a learner, and modeling techniques that are being considered in the learner model when developing an adaptive intelligent tutoring system.

Over the past years, learner modeling has been a matured topic and can encourage technology to be utilized for the implementation of adaptivity and personalization in ITS. This study presents a literature review of learner modeling form the year 2002 to 2017. This critical review also presents the methods that are to be used by a learner model to give an adaptive instruction to the learner. In addition to that, the study can also help pedagogy developers and researchers, who wish to integrate learner modeling in ITS. Furthermore, this study is significant to determine the possible techniques of learner modeling that were used in adaptive intelligent tutoring systems, which will be a guide to the future direction in the field. A lot of

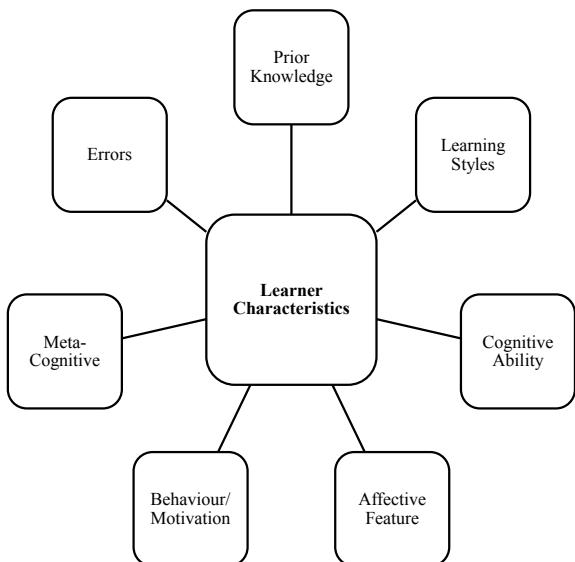
learner characteristics that have been used in learner modeling and the contribution to the development of learner models are reviewed here. In addition to that the study is expected to contribute to designing and developing of an intelligent and efficient learning system for future studies.

4 Learner Characteristics Model

Every learner learns differently and desires to learn in his/her individual way based on their needs, preferences, and characteristics like personal interest, skills, prior knowledge, learning styles, and cognitive, meta-cognitive abilities. Therefore, an essential phase of developing an adaptive learner model is to determine the suitable characteristics of the learner. To build a learner model, some questions have to be answered as follows: “Which are the important aspects of learner that would it be a good idea for us to model in a computerized educational system?” [15]. In order to implement the learner model, it needs to consider the psychological and nonpsychological parameter of the learner, which are sometimes called as static and dynamic characteristics of the learner. The common learner characteristics are shown in Fig. 1.

The concept of static and dynamic characteristics of the learner [16]. Static characteristics of the learners are defined by the personal details such as, name, email, language, and age, which are collected in pre-tutoring phase through the questionnaire and remain unchanged during the tutoring session. The dynamic characteristics of the learner include his/her prior domain knowledge level, learning styles, affective state factor [17], cognitive, and meta-cognitive factor which are collected during the

Fig. 1 Learner characteristics model



ongoing tutoring sessions [18]. These are also termed as performance features and may be a combination of static or dynamic learning data, and updates during the tutoring session to make learning adaptive.

In this manner, considering the appropriate dynamic characteristics of the learner is a challenge, which helps for system adaptation to every learner's needs and preferences. The extent of learner prior knowledge on the domain or current domain knowledge represent the learner domain knowledge level. Learning style is the most important characteristics of the learner for determining the individual preference of learner and play a crucial role in adaptive e-learning systems [19, 20].

Learning style describes the mixture of learner characteristics/attributes such as psychological, generic, emotive, and cognitive attributes works as a stable measure to help how a learner store, process, and responds to the learning conditions. Learning style has been classified into five learning dimensions such as Active and Reflective, Verbal and Visual, Sequential and Global, Inductive and Deductive, Sensory and Intuitive. As per a few investigations, adaptive frameworks in light of learning style are more gainful, make higher fulfillment level, decrease learning time, and increment the scholastic accomplishment of learner [21–23].

Considering the classroom teaching, to improve the learning goals and to make learning effective, human tutors observe the emotional state of the learner. Intelligent tutoring system should make itself adaptable by interpreting the emotional traits of learners by considering the learner needs and preferences [24]. Therefore, affective states should be considered for constructing the learner model. According to several researchers [25, 26], emotional-affective states can be happiness, frustrated, surprise, fear, disgusting, sadness, neutral, boring, focused, and confused. Some of these emotions like frustration or boring lead to a determination of the degree of engagement or off-task behavior of the learner. Modeling the emotional state forms a basis of the learning modeling in the adaptive intelligent tutoring system.

The important characteristics that should be considered in the learner modeling are learner cognitive traits/features such as the ability to understand, perception, attention, memory, ability to learn new information, error and misconception, learning new skills, ability to problem-solving, reasoning, and critical thinking, ability to processing thoughts, concentration and decision making abilities. However, apart from cognitive abilities, there is also a need to critically assess the learner knowledge to decide the learning contents to be offered for learning [27]. Thus, for providing adaptivity and personalization in a tutoring system, we must contemplate the meta-cognitive aspect of the learner in learner modeling. Meta-cognitive aspect/skills refer to learners evaluating themselves, memory monitoring, regulation, auto-consciousness, or self-awareness. Therefore, the metacognition skills help the learner to deal with his/her ability and critical thinking.

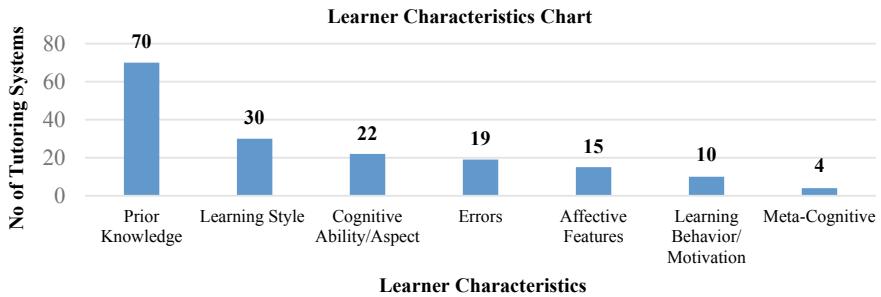


Fig. 2 Utilization of learner characteristics

5 Results and Discussion

The methodology of development of an adaptable ITS through the development of a learner model adapted through learner characteristics is presented in this section. This section presents the components and flow of proposed learner model framework to construct an adaptable learner model based on the learner characteristics shown in Fig. 2 (Table 1).

6 Conclusion

This paper was focused on the contribution of learner characteristics in the development of an adaptive learner model, which is the most important component of AITS. The learner characteristics, data, and information presented by a learner model were analyzed from 2002 to 2017. It took a gander at its history, inspiration and additionally investigated some outstanding ITSs. The analysis can help researchers in selecting the appropriate learner characteristics to develop an adaptive learner model. In addition to the findings, researchers can find which characteristics of the learner should be used most for implementation of adaptive learner model in ITS. The study reveals that the most important characteristics of the learner were prior knowledge, learning style, and cognitive skills, which were used in the development of the learner model. In the recent past, the affective features, such emotion and learning behavior of a learner was used in the adaptive learner modeling capabilities .

Table 1 Summary of common learner characteristics used in adaptive/personalized educational systems

S.N	Learner Characteristics	Frequency (f)	Intelligent Tutoring System (ITS)/E-learning/Personalized Learning System/Web-Based Educational System/Computerized Tutor/Adaptive System
1	Prior Knowledge	70	E-Teacher, ADAPTAPlan, DeLC, AMPLIA, English ABLE, Auto-Colleague, Chrysafidi and Virvou [28–30]; DesignFirstITS, InfoMap, Collect-UML, VIRGO, Web-IT, ICICLE, KER MIT, Andes, Millán and Pérez de la Cruz [31]; F-SMILE, INSPIRE, Web-PTV, Surjono and Malby [32]; SOQT-Web, TADV, Xu et al. [33], Leco-EAD, ACE, F-CBR-DHTC, XAHM, WILEDS, METADOC MEDEA, ISIS-TUTOR Intermediate Actor, Personal Reader, Baker et al. [34], AVANTI, ANANTOM-TUTOR, Stathacopoulou et al. [36–38]; INQPRO, DEPHS, SonITS, MAEVIF, IWT, AdpatErrEx, LS-Plan, Wang et al. [39]; J-LATTE, INCOM, CLT, ABM, Baschera and Gross [40, 41]; POOLE III, Baker et al. [35]; Jia et al. [42–45]; Kassim et al. [54–56]
2	Learning Style	30	GIAS, LS-Plan, Pranitasari et al. [38, 39]; WELSA, ABM, Al-Hmouz et al. [45]; Jia et al. [42]; Ayala and Sossa [46]; OPAL, IWT, MAEVIF, DEPHS, KBS HYPERBOOK, AHS-TS, INSPIRE, Surjono and Malby [32]; ACE, TADV, The personal Reader, Stathacopoulou et al. [36]; HYPERPLEX, XAHM, Salim and Haron [47–49]; DesignFirstITS, ADAPTAPlan, E-Teacher, InterBook
3	Cognitive Ability/Aspect	22	Andes, Xu et al. [33]; Web-PTV, F-CBR-DHTC, ICICLE, VIRGE, DeLC, Auto-Colleague, Kofod-Petersen et al. [50]; AMPLIA, ADAPTAPlan, GIAS, Wang et al. [39]; CLT, ABM, Milan et al. [43]; Jia et al. [42]; Ayala and Sossa [46]; TELEOS, DEPHS, AHS-TS, Al- Hmouz et al. [45];
4	Errors	19	Andes et al.; F-SMILE, Surjono and Malby [32], LeCo-EAD, Baker et al. [34]; Auto-Colleague, Chrysafidi and Virvou [28], KERMIT, InfoMap, J-LATTE, INCOM, Baschera and Gros [40]; Milan et al. [43, 44]; EER-Tutor, AdaptErrEx, AdaptErrEx, Kassim et al. [54]; Kumar [29, 30]; Le and Menzel [55]
5	Affective Features	15	Conati and MacLaren [51]; ABM, POOLE III, Millán et al. [43]; PlayPhysics, Wayang Outpost, Balkrishnan [44]; Crystal Island, MMT, Conati and Zhou [52]; VIRGE, Baker [48], Auto-Colleague, Kofod-Petersen et al. [50]; Alepis and Viryou [53]
6	Learning Behavior/Motivation	10	Centintas et al. [41]; Millán et al. [43]; PlayPhysics, OPAL, MMT, Baker [48]; VIRGE, Baker et al. [34], AdaptaWeb AHA
7	Meta-Cognitive	4	Andes, Millán et al. [43]; Wayang Outpost, IWT

References

1. Kumar, A., Singh, N., Ahuja, N.J.: Learner characteristics based learning style models classification and its implications on teaching. *Int. J. Pure Appl. Math. (IJPAM)* **118**(20), 175–184 (2017)
2. Kumar, A., Singh, N., Ahuja, N.J.: Learning styles based adaptive intelligent tutoring systems: Document analysis of articles published between 2001. and 2016. *Int. J. Cogn. Res. Sci. Eng. Educ. (IJCRSEE)* **5**(2), 83–98 (2017)
3. Bull, S., Greer, J., McCalla, G., Kettell, L., Bowes, J.: User modelling in i-help: what, why, when and how. In: International Conference on User Modeling, pp. 117–126. Springer, Berlin. https://doi.org/10.1007/3-540-44566-8_12 (2001)
4. Bull, S.: Supporting learning with open learner models. *Planning* **29**(14), 1 (2004)
5. Brusilovsky, P.: Adaptive hypermedia. *User Model. User-Adapt. Interact.* **11**(1–2), 87–110 (2001)
6. Brusilovsky, P., Millán, E.: User models for adaptive hypermedia and adaptive educational systems. In: *The Adaptive Web*, pp. 3–53. Springer, Berlin. https://doi.org/10.1007/978-3-540-72079-9_1 (2007)
7. Wenger, E.: Artificial intelligence and tutoring systems: computational and cognitive approaches to the communication of knowledge. Morgan Kaufman Publishers Inc., San Francisco (1987). ISBN 0-934613-26-5
8. Brusilovsky, P.: Student model centered architecture for intelligent learning environments. In: Proceedings of Fourth International Conference on User Modeling, Hyannis, MA, pp. 31–36, 15–19 Aug 1994. MITRE (1994)
9. Self, J., Lawler, R.W., Yazdani, M.: The application of machine learning to student modelling. In: *Artificial Intelligence and Education*, vol. 1, pp. 267–280. Ablex, Norwood, New Jersey. <https://doi.org/10.1007/BF00051826> (1987)
10. Self, J.A.: Realism in student modelling. In: Alvey-IKBS Research Workshop Tutoring Systems. University of Exeter
11. Self, J. (ed.): *Artificial Intelligence and Human Learning: Intelligent Computer-Aided Instruction*. Chapman and Hall, London (1988a)
12. Self, J.A.: Student models: what use are they? In: Ercoli, P., Lewis, R. (eds.) *Artificial Intelligence Tools in Education*, pp. 73–86. North-Holland, Amsterdam (1988b)
13. Gilmore, D., Self, J.: The application of machine learning to intelligent tutoring systems. In: *Artificial Intelligence and Human Learning, Intelligent Computer-Assisted Instruction*, pp. 179–196. Chapman and Hall, New York (1988)
14. Li, N., Cohen, W.W., Koedinger, K.R., Matsuda, N.: A machine learning approach for automatic student model discovery. In: *EDM*, pp. 31–40 (2011)
15. González, C., Burguillo, J. C., Llamas, M.: A qualitative comparison of techniques for student modeling in intelligent tutoring systems. In: *Frontiers in Education Conference*, 36th Annual, pp. 13–18 (2006)
16. Jeremić, Z., Jovanović, J., Gasević, D.: Student modeling and assessment in intelligent tutoring of software patterns. *Expert. Syst. Appl.* **39**(1), 210–222 (2012)
17. Graesser, A.C., D'Mello, S.: Moment-To-moment emotions during reading. *Read. Teach.* **66**(3), 238–242 (2012)
18. Essalmi, F., Ayed, L.J.B., Jemni, M., Graf, S.: A fully personalization strategy of e-learning scenarios. *Comput. Hum. Behav.* **26**(4), 581–591 (2010)
19. Ozyurt, O., Ozyurt, H.: Learning styles based individualized adaptive e-learning environments: content analysis of the articles published from 2005 to 2015. *Comput. Hum. Behav.* **52**, 349–358 (2015)
20. Truong, H.M.: Integrating learning styles and adaptive e-learning system: current developments, problems and opportunities. *Comput. Hum. Behav.* **55**, 1185–1193 (2016)
21. Tseng, J.C.R., Chu, H.C., Hwang, G.J., Tsai, C.C.: Development of an adaptive learning system with two sources of personalization information. *Comput. Educ.* **51**(2), 776–786 (2008)

22. Graf, S., Liu, T.C., Kinshuk. Analysis of learners' navigational behavior and their learning styles in an online course. *J. Comput. Assist. Learn.* **26**(2), 116–131 (2010)
23. Graf, S., Liu, T.C., Kinshuk, Chen, N.S., Yang, S.J.H.: Learning styles and cognitive traits—their relationship and its benefits in web-based educational systems. *Comput. Hum. Behav.* **25**(6), 1280–1289 (2009)
24. Katsionis, G., Virvou, M.: A cognitive theory for affective user modelling in a virtual reality educational game. In: 2004 IEEE International Conference on Systems, Man and Cybernetics, vol. 2, pp. 1209–1213. IEEE (2004)
25. Balakrishnan, A.: On modeling the affective effect on learning. In: International Workshop on Multi-disciplinary Trends in Artificial Intelligence, pp. 225–235. Springer, Berlin (2011)
26. Dorça, F.A., Lima, L.V., Fernandes, M.A., Lopes, C.R.: Comparing strategies for modeling students learning styles through reinforcement learning in adaptive and intelligent educational systems: An experimental analysis. *Expert Syst. Appl.* **40**(6), 2092–2101 (2013)
27. Mitrovic, A., Martin, B.: Evaluating the effects of open student models on learning. In: Proceedings of the 2nd International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems, pp. 296–305 (2006)
28. Chrysafiadi, K., Virvou, M.: Personalized teaching of a programming language over the web: Stereotypes and rule-based mechanisms. In: JCKBSE (pp. 484–492) (2008, June)
29. Kumar, A.N.: Using Enhanced Concept Map for Student Modeling in Programming Tutors. In FLAIRS Conference, pp. 527–532 (2006a)
30. Kumar, A.: A scalable solution for adaptive problem sequencing and its evaluation. In: International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems, pp. 161–171. Springer, Berlin, Heidelberg (2006b, June)
31. Millán, E., Pérez-De-La-Cruz, J.L.: A Bayesian diagnostic algorithm for student modeling and its evaluation. *User Model. User-Adap. Inter.* **12**(2–3), 281–330 (2002)
32. Surjono, H.D., Maltby, J.R.: Adaptive educational hypermedia based on multiple student characteristics. In International Conference on Web-Based Learning, pp. 442–449. Springer, Berlin, Heidelberg (2003, August)
33. Xu, D., Wang, H., Su, K.: Intelligent student profiling with fuzzy models. In Proceedings of the 35th Annual Hawaii International Conference on System Sciences, pp. 8–pp IEEE (2002, January)
34. Baker, R.S., Corbett, A.T., Koedinger, K.R., Wagner, A.Z.: Off-task behavior in the cognitive tutor classroom: when students game the system. In Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 383–390. ACM (2004, April)
35. Baker, R.S., Goldstein, A.B., Heffernan, N.T.: Detecting the moment of learning. In International Conference on Intelligent Tutoring Systems, pp. 25–34. Springer, Berlin, Heidelberg (2010, June)
36. Stathacopoulou, R., Magoulas, G.D., Grigoriadou, M., Samarakou, M.: Neuro-fuzzy knowledge processing in intelligent learning environments for improved student diagnosis. *Inf. Sci.* **170**(2–4), 273–307 (2005)
37. Goel, G., Lallé, S., Luengo, V.: Fuzzy logic representation for student modelling. In International Conference on Intelligent Tutoring Systems, pp. 428–433. Springer, Berlin, Heidelberg (2012, June)
38. Pramitasari, L., Hidayanto, A.N., Aminah, S., Krisnadhi, A.A., Ramadhanie, A.M.: Development of student model ontology for personalization in an e-learning system based on semantic web. In International Conference on Advanced Computer Science and Information Systems (ICACSIS09), Indonesia, December (pp. 7–8) (2009, December)
39. Wang, X., Yang, Y., Wen, X.: Study on blended learning approach for english teaching. In: 2009 IEEE International Conference on Systems, Man and Cybernetics, pp. 4641–4644. IEEE (2009, October)
40. Baschera, G.M., Gross, M.: Poisson-based inference for perturbation models in adaptive spelling training. *Int. J. Artif. Intell. Educ.* **20**(4), 333–360 (2010)
41. Cetintas, S., Si, L., Xin, Y.P.P., Hord, C.: Automatic detection of off-task behaviors in intelligent tutoring systems with machine learning techniques. *IEEE T Learn Technol.* **3**(3), 228–236 (2009)

42. Jia, B., Zhong, S., Zheng, T., Liu, Z.: The study and design of adaptive learning system based on fuzzy set theory. In *Transactions on edutainment IV*, pp. 1–11. Springer, Berlin, Heidelberg (2010)
43. Millán, E., Loboda, T., Pérez-De-La-Cruz, J.L.: Bayesian networks for student model engineering. *Comput. Educ.* **55**(4), 1663–1683 (2010)
44. Balakrishnan, A.: On modeling the affective effect on learning. In *Proceedings of the 5th international conference on multi-disciplinary trends in artificial intelligence*, pp. 225–235. Hyderabad, India (2011)
45. Al-Hmouz, A., Shen, J., Yan, J., Al-Hmouz, R.: Enhanced learner model for adaptive mobile learning. In *Proceedings of the 12th international conference on information integration and web-based applications & services*, pp. 783–786. ACM. (2010, November)
46. Ayala, A.P., Sossa, H.: Semantic representation and management of student models: an approach to adapt lecture sequencing to enhance learning. In *Mexican International Conference on Artificial Intelligence*, pp. 175–186. Springer, Berlin, Heidelberg (2010, November)
47. Salim, N., Haron, N.: The construction of fuzzy set and fuzzy rule for mixed approach in adaptive hypermedia learning system. In *International Conference on Technologies for E-Learning and Digital Entertainment*, pp. 183–187. Springer, Berlin, Heidelberg (2006, April)
48. Baker, R.S.: Modeling and understanding students' off-task behavior in intelligent tutoring systems. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 1059–1068. ACM (2007, April)
49. Carmona, C., Conejo, R.: A learner model in a distributed environment. In *International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*, pp. 353–359. Springer, Berlin, Heidelberg (2004, August)
50. Kofod-Petersen, A., Abbas, S., Gunhild, P., Bye, G., Kolås, L., Staupe, A.: Learning in an ambient intelligent environment towards modelling learners through stereotypes (2008)
51. Conati, C., Maclare, H.: Empirically building and evaluating a probabilistic model of user affect. *User Model. User-Adapt. Interact.* **19**(3), 267–303 (2009)
52. Conati, C., Zhou, X.: Modeling students' emotions from cognitive appraisal in educational games. In *International Conference on Intelligent Tutoring Systems*, pp. 944–954. Springer, Berlin, Heidelberg (2002, June)
53. Alepis, E., Virvou, M.: Automatic generation of emotions in tutoring agents for affective e-learning in medical education. *Expert Syst. Appl.* **38**(8), 9840–9847 (2011)
54. Kassim, A.A., Kazi, S.A., Ranganath, S. U. R. E. N. D. R. A.: A web-based intelligent learning environment for digital systems. *Int. J. Eng. Educ.* **20**(1), 13–23 (2004).
55. Le, N.T., Menzel, W.: Using weighted constraints to diagnose errors in logic programming—the case of an Ill-defined domain. *Int. J. Artif. Intell. Educ.* **19**(4), 381–400 (2009)
56. Martin, B.: Constraint-based student modeling: representing student knowledge. In *New Zealand Computer Science Research Students' Conference*, pp. 22–29 (1999)

Event-Triggered Based Synchronization of Linear Singularly Perturbed Systems



Nakul Kotibhaskar, Kritika Bansal and Pankaj Mukhija

Abstract This paper addresses the synchronization problem in a complex dynamical network where each node is considered a linear singularly perturbed system. A network is considered where the slow dynamics of one node may be coupled with fast dynamics of the other nodes and vice versa. Two-time scale separation of the overall network is shown and feedback control is designed using classical singular perturbation theory. Considering the limited network resources, an event-triggering mechanism is designed separately for the slow and fast states of the overall network, such that synchronization is achieved without Zeno behaviour.

Keywords Synchronization · Singular perturbation · Event-triggered · Complex dynamical network

1 Introduction

Various systems occurring in nature such as metabolic systems, telephone networks and the internet can be modelled using complex dynamical networks (CDN) (see [1] and references therein). An CDN is a set of interconnected nodes, in which each node can have its own certain dynamics. In many networks such as power systems and neural networks, the dynamics of these nodes have a two-time scale property, namely the ‘slow’ and the ‘fast’ dynamics. Such systems possessing both slow and fast dynamics are called singularly perturbed systems (SPS). In order to stabilize such systems, the control design is carried out by reducing the system into approximate lower order slow and fast states and then separately analysing them [2].

In the context of CDNs, the synchronization phenomena has been receiving much attention since the past few years and various techniques have been developed to address the synchronization problem for regular systems such as in [3–6]. Synchronization problem in networks of SPS has been explored in [7–10]. In [7–9], singularly perturbed complex networks (SPCN) have been considered that assume

N. Kotibhaskar (✉)
New Delhi, India

K. Bansal · P. Mukhija
National Institute of Technology Delhi, New Delhi, India

the slow dynamics of each node are only coupled to the slow dynamics of the other nodes and likewise for fast states. However in [10], such a coupling constraint is relaxed and sufficient conditions are obtained for synchronization of a linear SPCN with coupling delays. All the above works consider time-triggered control for solving the synchronization problem.

However, with the ever increasing size and complexity of CDNs, network constraints must be taken into consideration for efficient utilization of network resources. Event-triggered control (ETC) is the answer to such constrained communication and has a clear advantage over time-triggered control as displayed in [11] and references therein. In [12, 13], ETC methodology has been applied to achieve synchronization in CDNs. But again, ETC applied to a network of SPS is yet to gain attention and not much work has been done in this area. In [14], ETC has been applied to a special class of non-linear SPCN with coupling delays to achieve synchronization.

Inspired by the above discussion, we explore the problem of synchronization for a complex network of linear SPS using ETC. A composite feedback law is designed to achieve synchronization of the network possessing two time scale behaviour. A centralized event-triggering mechanism is proposed that determines the triggering instants independently for the overall slow and fast states of the network. The case where synchronization is achieved using ETC based only on feedback of slow states is covered. The following contributions are then presented as:

- Stability using ETC is explored for a network of linear SPS, where the inner coupling matrix has a general form, i.e. the slow states of one node may be coupled with fast states of the other nodes and vice versa.
- Synchronization using event-triggered feedback based only on the slow dynamics is achieved, while ensuring a positive minimum inter-event time (MIET) for both the slow and fast states.
- Proposed methodology is used to guarantee the stability of a single SPS.

2 Problem Formulation and Preliminaries

The standard model for a linear SPS is given by

$$\begin{cases} \dot{x}(t) = A_{11}x(t) + A_{12}z(t) + B_1u(t) \\ \epsilon\dot{z}(t) = A_{21}x(t) + A_{22}z(t) + B_2u(t) \end{cases} \quad (1)$$

where $x \in R^{n_x}$, $z \in R^{n_z}$, $u \in R^m$, $0 < \epsilon << 1$ being the singular perturbation parameter. The set of Eq. (1) can be written in compact form as

$$\dot{X}(t) = A_\epsilon X(t) + B_\epsilon u(t) \quad (2)$$

where $X = [x^T \ z^T]^T$, $A_\epsilon = \begin{bmatrix} A_{11} & A_{12} \\ \frac{A_{21}}{\epsilon} & \frac{A_{22}}{\epsilon} \end{bmatrix}$ and $B_\epsilon = \begin{bmatrix} B_1 \\ \frac{B_2}{\epsilon} \end{bmatrix}$.

Hence, the dynamics of the i th node in the complex network comprising of N identical nodes can be modelled as

$$\dot{X}_i(t) = A_\epsilon X_i(t) + c \sum_{j=1}^N G_{ij} M X_j(t) + B_\epsilon u_i(t), \quad i = 1, 2 \dots N \quad (3)$$

where $c \geq 0$ is the coupling strength, $M \in R^{(n_x+n_y) \times (n_x+n_y)}$ is a constant inner coupling matrix of the nodes, $G = (G_{ij})_{N \times N}$ is the outer coupling matrix of the nodes, where G_{ij} is defined as follows: if there is a connection between node i and node j ($j \neq i$), then $G_{ij} = G_{ji} = 1$; otherwise, $G_{ij} = G_{ji} = 0$, and the diagonal elements of matrix G are defined by

$$G_{ii} = - \sum_{j=1, j \neq i}^N G_{ij} = - \sum_{j=1, j \neq i}^N G_{ji}, \quad i = 1, 2 \dots N \quad (4)$$

Synchronization error is denoted by $q_i(t) = X_i(t) - S(t)$, where $S(t)$ is the state trajectory of the solution of unforced isolated node given by $\dot{X}(t) = A_\epsilon X(t)$. Then, the error dynamics of the i th node of the network defined by (3) is given by

$$\dot{q}_i(t) = A_\epsilon q_i(t) + c \sum_{j=1}^N G_{ij} M q_j(t) + B_\epsilon u_i(t), \quad i = 1 \dots N \quad (5)$$

Using the Kronecker product, Eq. (5) can be written in compact form

$$\dot{q}(t) = (I_N \otimes A_\epsilon)q(t) + c(G \otimes M)q(t) + (I_N \otimes B_\epsilon)U(t) \quad (6)$$

where $q(t) = [q_1(t)^T \ q_2(t)^T \ \dots \ q_N(t)^T]^T$ and $U = [u_1(t)^T \ u_2(t)^T \ \dots \ u_N(t)^T]^T$.

In order to treat the fast and slow dynamics of the overall network independently, rearrangement of the state vectors is done using a product of series of row switching transformation matrices given by $\hat{T} = T_1 T_2 \dots T_N \in R^{N(n_x+n_z) \times N(n_x+n_z)}$. Then, using the properties of elementary row switching transformations it can be seen that $\hat{T} = \hat{T}^{-1}$ and

$$\hat{T} \begin{bmatrix} q_{x_1}(t) \\ q_{z_1}(t) \\ q_{x_2}(t) \\ q_{z_2}(t) \\ \vdots \\ \vdots \\ q_{x_N}(t) \\ q_{z_N}(t) \end{bmatrix} = \begin{bmatrix} q_{x_1}(t) \\ q_{x_2}(t) \\ \vdots \\ q_{x_N}(t) \\ q_{z_1}(t) \\ q_{z_2}(t) \\ \vdots \\ q_{z_N}(t) \end{bmatrix}, \quad \hat{T} \begin{bmatrix} B_1 u_1(t) \\ \frac{B_2}{\epsilon} u_1(t) \\ B_1 u_2(t) \\ \frac{B_2}{\epsilon} u_2(t) \\ \vdots \\ B_1 u_N(t) \\ \frac{B_2}{\epsilon} u_N(t) \end{bmatrix} = \begin{bmatrix} B_1 u_1(t) \\ B_1 u_2(t) \\ \vdots \\ B_1 u_N(t) \\ \frac{B_2}{\epsilon} u_1(t) \\ \frac{B_2}{\epsilon} u_2(t) \\ \vdots \\ \frac{B_2}{\epsilon} u_N(t) \end{bmatrix}$$

Similarly,

$$\hat{T}(I_N \otimes A_\epsilon)\hat{T}^{-1} = \begin{bmatrix} I_N \otimes A_{11} & I_N \otimes A_{12} \\ I_N \otimes \frac{A_{12}}{\epsilon} & I_N \otimes \frac{A_{22}}{\epsilon} \end{bmatrix}, \quad \hat{T}(G \otimes M)\hat{T}^{-1} = \begin{bmatrix} G \otimes M_{11} & G \otimes M_{12} \\ G \otimes M_{21} & G \otimes M_{22} \end{bmatrix}$$

The system then becomes

$$\hat{T}\dot{q}(t) = \hat{T}(I_N \otimes A_\epsilon)\hat{T}^{-1}\hat{T}q(t) + c\hat{T}(G \otimes M)\hat{T}^{-1}\hat{T}q(t) + \hat{T}(I_N \otimes B_\epsilon)U(t) \quad (7)$$

Further, in order to write the above equation in the form of a general SPS, both sides of the equation are pre-multiplied by $\text{diag}\{I_{Nn_x}, \epsilon I_{Nn_z}\}$. Finally, the error dynamical system of the complex network in compact form is given by

$$E_\epsilon \dot{\bar{q}}(t) = \bar{A}\bar{q}(t) + cE_\epsilon \bar{G}\bar{q}(t) + \bar{B}U(t) \quad (8)$$

$$\text{where } \bar{q} = [q_x^T \ q_z^T]^T \text{ and } E_\epsilon = \begin{bmatrix} I_{Nn_x} & 0 \\ 0 & \epsilon I_{Nn_z} \end{bmatrix}, \quad \bar{A} = \begin{bmatrix} I_N \otimes A_{11} & I_N \otimes A_{12} \\ I_N \otimes A_{21} & I_N \otimes A_{22} \end{bmatrix}, \quad \bar{G} = \begin{bmatrix} G \otimes M_{11} & G \otimes M_{12} \\ G \otimes M_{21} & G \otimes M_{22} \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} I_N \otimes B_1 \\ I_N \otimes B_2 \end{bmatrix}$$

Definition 1 The complex network (3) is said to achieve synchronization if there exists a constant $0 < \epsilon^* << 1$ such that $X_1(t) = X_2(t) = \dots = X_N(t) = S(t)$, for all $\epsilon \in (0, \epsilon^*]$. Hence, $q(t) = 0$ (equivalently, $\bar{q}(t) = 0$) as $t \rightarrow \infty$. Therefore, the complex network is asymptotically synchronized if the error dynamical system (8) is asymptotically stable at the origin.

3 Main Results

3.1 Controller Design

The overall system is approximated using a slow and a fast model and Kronecker product properties (see [15]) have been utilized to accomplish it. Assuming A_{22}^{-1} exists, letting $\epsilon = 0$ for the error system (8), the reduced slow system for the network is given by

$$\dot{q}_{x_s}(t) = \bar{A}_0 q_{x_s}(t) + c\bar{G}_0 q_{x_s}(t) + \bar{B}_0 U_s(t) \quad (9)$$

$$q_{z_s}(t) = -(I_N \otimes A_{22}^{-1} A_{21}) q_{x_s}(t) - (I_N \otimes A_{22}^{-1} B_2) U_s(t) \quad (10)$$

where $q_{x_s}(t)$, $q_{z_s}(t)$, $U_s(t)$ are slow parts of the corresponding state vectors and

$$\bar{A}_0 = I_N \otimes (A_{11} - A_{12} A_{22}^{-1} A_{21}), \quad \bar{G}_0 = G \otimes (M_{11} - M_{12} A_{22}^{-1} A_{21})$$

$$\bar{B}_0 = (I_N \otimes B_1) - (I_N \otimes A_{12} A_{22}^{-1} B_2) - c(G \otimes M_{12} A_{22}^{-1} B_2)$$

Similarly, the boundary layer fast system for the network is given by

$$\epsilon \dot{q}_{z_f}(t) = (I_N \otimes A_{22})q_{z_f}(t) + (I_N \otimes B_2)U_f(t) \quad (11)$$

Control law is then designed independently for the slow and fast dynamics as $U_s(t) = K_0 q_{x_s}(t)$, $U_f = K_2 q_{z_f}(t)$, where $K_0 = \text{diag}\{K_{01}, K_{02} \dots K_{0N}\}$, $K_2 = \text{diag}\{K_{21}, K_{22} \dots K_{2N}\}$. Composite control is then designed as: $U_s + U_f = K_0 q_{x_s} + K_2 q_{z_f}$. For a realizable feedback, q_{x_s} is replaced by q_x and q_{z_f} is replaced by $q_z - q_{z_s}$ such that

$$U(t) = K_1 q_x(t) + K_2 q_z(t) \quad (12)$$

where $K_1 = (I_{Nm} + K_2(I_N \otimes A_{22}^{-1}B_2))K_0 + K_2(I_N \otimes A_{22}^{-1}A_{21})$. Then from [2], feedback gains K_1 and K_2 can be designed to render error dynamical system (8) asymptotically stable for an $\epsilon^* \geq 0$. Hence, by Definition 1, the control law (12) asymptotically synchronizes the complex network (3). Also from [2], if A_{22} is Hurwitz, K_0 can be designed such that control law $U = K_0 q_x(t)$ asymptotically stabilizes error dynamical system (8) leading to asymptotic synchronization of complex network (3).

3.2 Event-Triggered Control

In this section, an ETC strategy is proposed to reduce the usage of network resources for achieving synchronization in CDN (3). Let $\{t_k^x\}$ and $\{t_k^z\}$ for $k \in \{0, 1, 2 \dots\}$ be the triggering instants for the combined slow and combined fast states, respectively. Therefore, the input $U(t)$ then takes the form of

$$U(t) = K_1 \hat{q}_x(t) + K_2 \hat{q}_z(t) \quad (13)$$

where $\hat{q}_x(t)$ and $\hat{q}_z(t)$ represent the latest transmitted state of the slow and fast states respectively. Then, the error induced due to sampling is defined as

$$e_x(t) = \hat{q}_x(t) - q_x(t), \quad e_z(t) = \hat{q}_z(t) - q_z(t) \quad (14)$$

System (8) can then be rewritten considering (13) and (14) in the manner

$$E_\epsilon \dot{\bar{q}}(t) = (\bar{A} + c E_\epsilon \bar{G} + \bar{B} K) \bar{q}(t) + \bar{B} K e(t) \quad (15)$$

where $K = [K_1 \ K_2]$, $e = [e_x^T \ e_z^T]^T$.

Now, the ETM for the network is proposed as

$$\begin{aligned} t_{k+1}^x &= \inf\{t : t > t_k^x \mid f(q_x, e_x, t) > 0\} \\ t_{k+1}^z &= \inf\{t : t > t_k^z \mid f(q_z, e_z, t) > 0\} \end{aligned} \quad (16)$$

for $f(q_a, e_a, t) = ||e_a(t)|| - \sqrt{\sigma^2 ||q_a(t)||^2 + \beta^2 e^{-2\alpha t}}$, $\sigma \geq 0$, $\beta \geq 0$ and $\alpha > 0$ are design parameters and $t_0^x = t_0^z = 0$. This triggering rule was inspired from a similar ETM used in [13]. It should be noted that both σ and β cannot be zero simultaneously. The event controlled strategy is implemented such that at every transmission instant t_k^a , $\hat{q}_a(t_k^a)$ is updated, i.e. $\hat{q}_a(t_k^a) = q_a(t_k^a)$. This implies that $e_a^+(t_k^a) = 0$.

Theorem 1 Consider the system (15), event-triggering mechanism (16) with controller gains given by (12). Then, for $0 < \epsilon \leq \epsilon^*$ there exists matrices P , Q and constant δ satisfying Eqs. (17) and (18), such that the network (3) is asymptotically synchronized. Also, there exists a bound on MIET for both the slow and fast states.

$$\begin{aligned} (\bar{A} + cE_\epsilon\bar{G} + \bar{B}K)^T P + P^T(\bar{A} + cE_\epsilon\bar{G} + \bar{B}K) &= -Q \\ (E_\epsilon P) > 0, \quad (Q = Q^T) &> 0 \end{aligned} \quad (17)$$

$$-\lambda_{min}(Q) + 2\sigma\gamma = -\delta, \quad \delta \geq 0 \quad (18)$$

where $\gamma = ||P\bar{B}K||$, σ has been defined in ETM (16).

Proof Let the Lyapunov function be

$$V(\bar{q}(t)) = \bar{q}(t)^T E_\epsilon P \bar{q}(t) \quad (19)$$

Let the feedback gain matrices in Eq. (12) been designed to render the error dynamical system (8) asymptotically stable, hence $E_\epsilon^{-1}(\bar{A} + cE_\epsilon\bar{G} + \bar{B}K)$ is Hurwitz. Then, calculating the derivative of the above Lyapunov function along the trajectory of the system (15), gives $\dot{V} = -\bar{q}^T Q \bar{q} + 2\bar{q}^T P \bar{B} K e$. Since, a positive definite matrix Q satisfies : $\lambda_{min}(Q)||\bar{q}||^2 \leq \bar{q}^T Q \bar{q} \leq \lambda_{max}(Q)||\bar{q}||^2$. Then, $\dot{V} \leq -\lambda_{min}(Q)||\bar{q}||^2 + 2||\bar{q}||\gamma||e||$ where γ has been defined above. As the ETM enforces the condition $||e_a(t)|| < \sqrt{\sigma^2 ||q_a(t)||^2 + \beta^2 e^{-2\alpha t}}$, therefore, $||e||^2 \leq \sigma^2 ||\bar{q}(t)||^2 + 2\beta^2 e^{-2\alpha t} \leq (\sigma||\bar{q}(t)|| + \sqrt{2}\beta e^{-\alpha t})^2$. Hence, the overall sampling induced error (i.e. $e(t)$) is bounded as $||e|| \leq (\sigma||\bar{q}(t)|| + \sqrt{2}\beta e^{-\alpha t})$. For a value of σ satisfying Eq. (18), it can be shown in a similar way as done in [13] that

$$\begin{aligned} \dot{V} &= -(\delta - \delta_1)||\bar{q}||^2 - ||\bar{q}||(\delta_1||\bar{q}|| - 2\sqrt{2}\gamma\beta e^{-\alpha t}) \\ \dot{V} &\leq -(\delta - \delta_1)||\bar{q}||^2, \text{ when } ||\bar{q}|| \geq \zeta e^{-\alpha t} \end{aligned}$$

where $0 < \delta_1 < \delta$, $\zeta = \frac{2\sqrt{2}\gamma\beta}{\delta_1}$. Then, letting $g(t) = \zeta e^{-\alpha t}$, which satisfies $\lim_{t \rightarrow \infty} g(t) = 0$. Then, using lemma 3 in [13], the system (15) is asymptotically stable at the origin and hence by Definition 1 the complex network (3) is asymptotically synchronized.

In order to show that there exists a bound on MIET, the system (15) is decoupled in exact slow and fast states using the transformation (see [2]).

$$\begin{bmatrix} p_x(t) \\ p_z(t) \end{bmatrix} = \begin{bmatrix} I_{Nn_x} - \epsilon HL & -\epsilon H \\ L & I_{Nn_z} \end{bmatrix} \begin{bmatrix} q_x(t) \\ q_z(t) \end{bmatrix} \quad (20)$$

where p_x , p_z represent the exact slow and fast states, respectively, and matrices L,H satisfy the following equations:

$$\Lambda_{21} - \Lambda_{22}L + \epsilon L\Lambda_{11} - \epsilon L\Lambda_{12}L = 0 \quad (21)$$

$$\Lambda_{12} - H\Lambda_{22} + \epsilon\Lambda_{11}H - \epsilon\Lambda_{12}LH - \epsilon HL\Lambda_{12} = 0 \quad (22)$$

such that $\Lambda = \bar{A} + cE_\epsilon\bar{G} + \bar{B}K$. Hence, $\Lambda_{1j} = \bar{A}_{1j} + c\bar{G}_{1j} + \bar{B}_1K_j$ and $\Lambda_{2j} = \bar{A}_{2j} + c\bar{G}_{2j} + \bar{B}_2K_j$ where $\bar{A}_{ij} = I_N \otimes A_{ij}$, $\bar{G}_{ij} = G \otimes M_{ij}$, $\bar{B}_i = I_N \otimes B_i$. Now, (15) becomes

$$\begin{bmatrix} \dot{p}_x(t) \\ \epsilon\dot{p}_z(t) \end{bmatrix} = \begin{bmatrix} A_s + B_s K_s & 0 \\ 0 & A_f + B_f K_2 \end{bmatrix} \begin{bmatrix} p_x(t) \\ p_z(t) \end{bmatrix} + \begin{bmatrix} B_s \\ B_f \end{bmatrix} Ke(t) \quad (23)$$

where,

$$\begin{aligned} A_s &\triangleq \bar{A}_0 + c\bar{G}_0 - \epsilon(\bar{A}_{12} + c\bar{G}_{12})\bar{A}_{22}^{-1}[c\bar{G}_{21} - c\bar{G}_{22}L + L(\bar{A}_{11} + c\bar{G}_{11}) - \\ &\quad L(\bar{A}_{12} + c\bar{G}_{12})L], \\ B_s &\triangleq \bar{B}_0 - \epsilon(\bar{A}_{12} + c\bar{G}_{12})\bar{A}_{22}^{-1}L\bar{B}_1, \\ A_f &\triangleq \bar{A}_{22} + \epsilon[c\bar{G}_{22} + L(\bar{A}_{12} + c\bar{G}_{12})], \quad B_f \triangleq \bar{B}_2 + \epsilon L\bar{B}_1, \quad K_s \triangleq K_1 - K_2L \end{aligned} \quad (24)$$

Using $O(\epsilon)$ approximations,

$$\begin{aligned} L &= \bar{A}_{22}^{-1}(\bar{A}_{21} + \bar{B}_2K_0) + O(\epsilon) \\ A_s + B_s K_s &= \bar{A}_0 + c\bar{G}_0 + \bar{B}_0K_0 + O(\epsilon) = \bar{\Lambda}_0 + O(\epsilon) \\ A_f + B_f K_f &= \bar{A}_{22} + \bar{B}_2K_2 + O(\epsilon) = \bar{\Lambda}_{22} + O(\epsilon) \end{aligned} \quad (25)$$

Using the above approximation, (23) can be equivalently written as

$$\begin{bmatrix} \dot{p}_x(t) \\ \dot{p}_z(t) \end{bmatrix} = A_D \begin{bmatrix} p_x(t) \\ p_z(t) \end{bmatrix} + B_D Ke(t) \quad (26)$$

where

$$A_D = \begin{bmatrix} \bar{\Lambda}_0 & 0 \\ 0 & \frac{\bar{\Lambda}_{22}}{\epsilon} \end{bmatrix} + O(\epsilon); \quad B_D = \begin{bmatrix} B_0 \\ \frac{B_2}{\epsilon} \end{bmatrix} + O(\epsilon) \quad (27)$$

For $t \geq 0$, q_x and q_z can be approximated using (20) as

$$q_x = p_x + O(\epsilon), \quad q_z = p_z - Lp_x + O(\epsilon) \quad (28)$$

First, consider the following triggering rule:

$$t_{k+1}^a = \inf\{t : t > t_k^a \mid \|e_a\| > \beta e^{-\alpha t}\}$$

Comparing Eqs. (9), (11) with (25), and using that $\bar{\Lambda}_0, \bar{\Lambda}_{22}$ are designed Hurwitz, hence $\|e^{\bar{\Lambda}_0 t}\| \leq \eta_1 e^{-\alpha_1 t}$, $\|e^{\bar{\Lambda}_{22} t}\| \leq \eta_2 e^{-\alpha_2 t}$ for positive constants $\eta_1, \eta_2, \alpha_1, \alpha_2$. Then, following in a similar way as case 1 from [16], it is obtained that the MIET for the slow states $(t_{k+1}^x - t_k^x)$ is lower bounded by τ given by equation

$$(k_1 + k_2)\tau = \beta e^{-\alpha\tau}, \quad (29)$$

is strictly positive, where

$$k_1 = \|\bar{\Lambda}_0\| \eta_1 \|p_x^0\|, \quad k_2 = \|\bar{B}_0 K\| \sqrt{2} \beta \left(1 + \frac{\eta_1 \|\bar{\Lambda}_0\|}{\alpha_1 - \alpha}\right) \quad (30)$$

Similarly, for the fast states the MIET for initial boundary period $[0, t_1]$ is lower bounded by τ' given by equation

$$(k'_1 + k'_2)\tau' = \beta e^{-\alpha\tau'} \quad (31)$$

which is strictly positive, where

$$k'_1 = \frac{\|\bar{\Lambda}_{22}\| \eta_2 \|p_z^0\|}{\epsilon}, \quad k'_2 = \frac{\|\bar{B}_2 K\| \sqrt{2} \beta}{\epsilon} \left(1 + \frac{\eta_2 \|\bar{\Lambda}_{22}\|}{\alpha_2 - \epsilon\alpha}\right) \quad (32)$$

For $t > t_1$, MIET is given by $\frac{\tau}{\|L\|}$, where τ is solution of (29). As for all $\sigma \geq 0$, $\sqrt{\sigma^2 \|q_a(t)\|^2 + \beta^2 e^{-2\alpha t}} \geq \beta e^{-\alpha t}$ is always satisfied, there also exists a positive lower bound on the MIET for the triggering rule (16) for both slow and fast states that are greater than those obtained in Eqs. (29) and (31).

Remark 1 From Sect. 3.1, control law $U = K_0 q_x(t)$ can asymptotically synchronize complex network (3) when A_{22} is Hurwitz. Applying event-triggered control with ETM (16) over such feedback law may not guarantee a positive lower bound for the MIET of the fast states as shown in [17]. As proposed in the same paper, by adding a dead zone in ETM, the network can achieve semi-global practical synchronization.

Remark 2 The feedback form considered in this paper enables us in analysing the stability of a single SPS. Also, separate ETM design for slow and fast dynamics helps in seamlessly reducing the problem to a single system where both the dynamics can be controlled independently and asynchronously as in [16].

Corollary 1 Consider the system (1), triggering mechanism (16) with $f(x, e_x, t)$ and $f(z, e_z, t)$ for slow and fast states, respectively. Then, for $0 < \epsilon \leq \epsilon^{**}$, there exists matrices S, W and constant δ satisfying Eqs. (33), (34) such that system (1) is asymptotically stable. Also, there exists a bound on MIET for both the slow and fast states.

$$\begin{aligned} (A + BK)^T S + S^T (A + BK) &= -W \\ (\tilde{E}_\epsilon S) &> 0; \quad (W = W^T) > 0 \end{aligned} \quad (33)$$

$$-\lambda_{\min}(W) + 2\sigma\gamma = -\delta, \quad \delta \geq 0 \quad (34)$$

where $\gamma = \|SBK\|$.

Proof It can be seen that when $N=1$ (i.e. $G=0$), synchronization problem for the complex dynamical network is reduced to stability problem for a single system (1), which can be written in the form of $\tilde{E}_\epsilon \dot{X} = (A + BK)X + BKe$. Using a Lyapunov function as $V(X) = X^T \tilde{E}_\epsilon S X$, the proof can be obtained in a similar way as done in Theorem 1.

4 Illustrative Example

In this section, simulation results obtained using MATLAB are presented to validate the efficacy of the above theoretical analysis. The first example deals with the synchronization problem. While the second example deals with the stability of a single system.

Example 1 A network consisting of three nodes is considered with $\epsilon = 0.01$, $n_x = 2$, $n_z = 2$, $c = 0.5$.

$$A_{11} = \begin{bmatrix} 2.5 & 6 \\ -2 & 2 \end{bmatrix}, A_{12} = \begin{bmatrix} 2 & 3 \\ 0 & -2 \end{bmatrix}, A_{21} = \begin{bmatrix} 0.5 & 2 \\ -1 & 1 \end{bmatrix}, A_{22} = \begin{bmatrix} -2 & 1 \\ 0 & -1 \end{bmatrix}, B_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, B_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$M = \begin{bmatrix} 3 & -0.50 & -0.81 & 0.7 \\ -0.5 & 2 & 0 & 0 \\ 1 & 0.6 & 0.71 & -0.2 \\ 0.3 & -0.1 & 1.11 & 0.325 \end{bmatrix}, G = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}$$

Since A_{22} is Hurwitz, using results of Sect. 3.1, feedback gains are obtained as $K_{01} = K_{02} = K_{03} = [-0.4794 \quad 0.2152]$. The following Remark 2, ETM is chosen as $t_{k+1}^x = \inf\{t : t > t_k^x \mid ||e_a(t)|| - \max\{\sqrt{\sigma^2||q_a(t)||^2 + \beta^2 e^{-2\alpha t}}, \rho\} > 0\}$ with $\sigma = 0.11$, $\rho = 0.001$, $\beta = 0$. As the feedback is based only on slow states, no event is triggered for fast states and the MIET for slow states is obtained as 0.066 with nine transmission instants as shown in Fig. 3. Figures 1 and 2 display the synchronization errors for the slow and fast states, respectively. Hence, the proposed feedback controller and ETM synchronizes the given network.

Example 2 Here, the example from [16] is taken and stability analysis is done using the ETM (16). Same values of feedback gain, sampling period, initial conditions are used and it can be seen from Table 1 that similar results are obtained using the proposed ETM with values $\sigma^2 = 0.16$, $\beta^2 = 0.04$ and $\alpha = 0.09$.

The state trajectories are plotted in Fig. 4 while Fig. 5 displays the triggering instants.

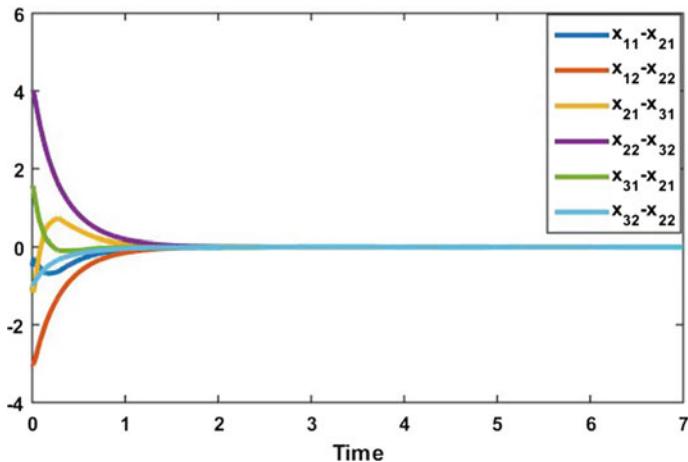


Fig. 1 Synchronization error for slow states

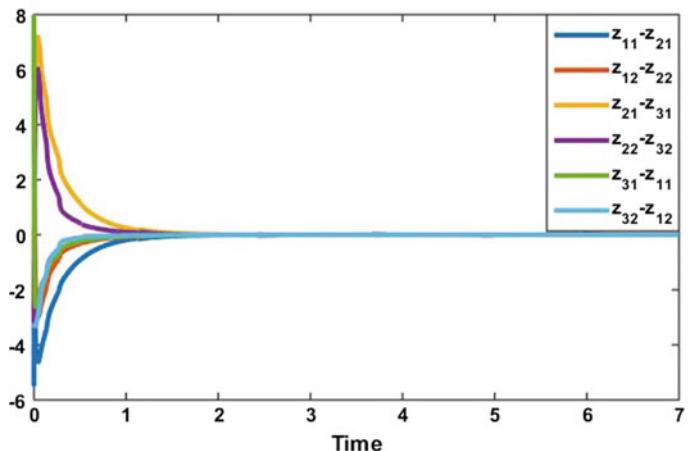


Fig. 2 Synchronization errors for fast states

5 Conclusion

Synchronization problem for networks of linear singularly perturbed systems using event-triggered control has been studied. Classical singular perturbation theory has been used for efficient feedback control design. Owing to the two-time scale property, an event-triggering strategy is proposed with independent triggering mechanisms for the overall slow and fast dynamics of the network. Asymptotic synchronization is achieved for the complex network with the proposed strategy without the occurrence

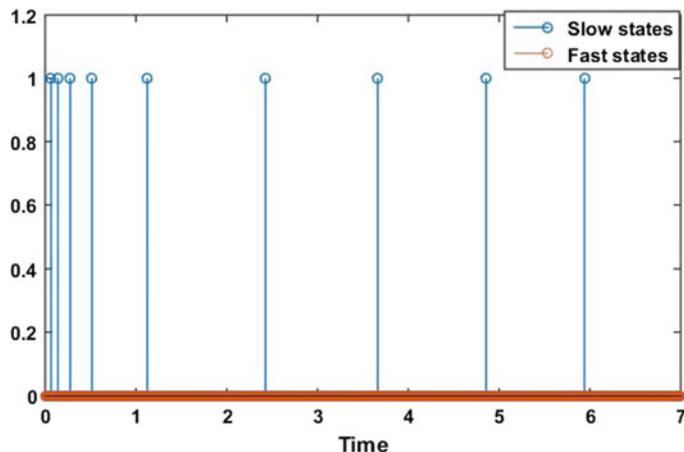


Fig. 3 Triggering instants

Table 1 Triggering events and inter-event times

	Slow states	Fast states
Number of events	8	11
Minimum inter-event time	1.6989	0.0543
Average inter-event time	2.2493	1.6857

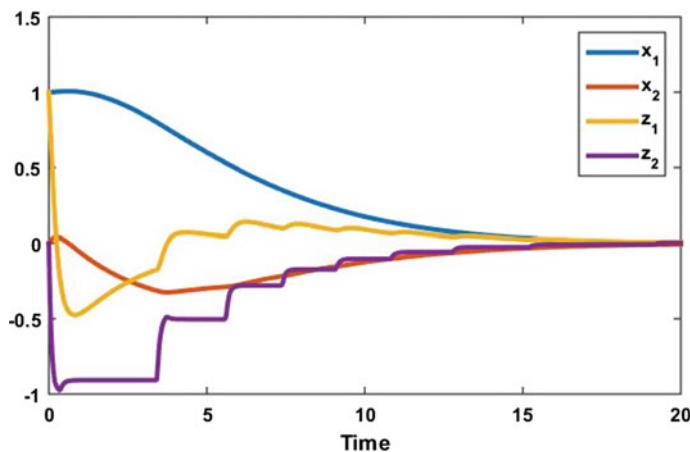


Fig. 4 State trajectories

of Zeno behaviour. The proposed methodology is used to guarantee the stability of a single singularly perturbed system. A numerical example is then used to show practical synchronization using event-triggered feedback based only on slow states.

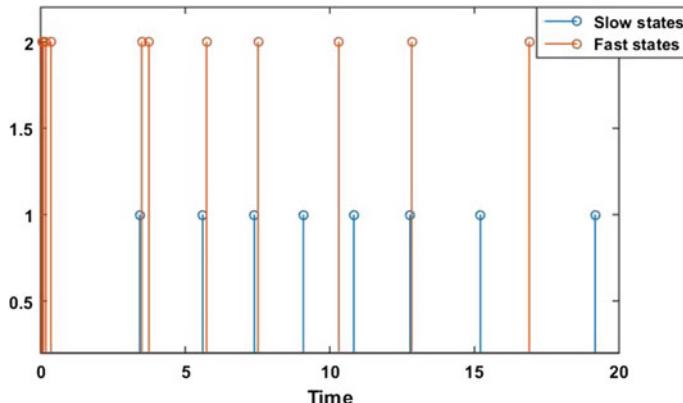


Fig. 5 Triggering instants

References

1. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. *Nature* **393**, 440 EP (1998)
2. Kokotovic, P., Khalil, H.K., O'Reilly, J.: Singular perturbation methods in control: analysis and design. SIAM (1999)
3. Wang, X.F., Chen, G.: Synchronization in small-world dynamical networks. *Int. J. Bifurc. Chaos* (2002)
4. Tu, L., Lu, J.A.: Delay-dependent synchronization in general complex delayed dynamical networks. *Comput. Math. Appl.* **57**(1), 28–36 (2009)
5. Li, N., Zhang, Y., Hu, J., Nie, Z.: Synchronization for general complex dynamical networks with sampled-data. *Neurocomputing* **74**(5), 805–811 (2011)
6. Xu, M., Wang, J.L., Huang, Y.L., Wei, P.C., Wang, S.X.: Pinning synchronization of complex dynamical networks with and without time-varying delay. *Neurocomputing* **266**, 263–273 (2017)
7. Rakkiyappan, R., Sivaranjani, K.: Sampled-data synchronization and state estimation for non-linear singularly perturbed complex networks with time-delays. *Nonlinear Dyn.* **84**(3), 1623–1636 (2016)
8. Cai, C., Wang, Z., Xu, J., Alsaedi, A.: Decomposition approach to exponential synchronisation for a class of non-linear singularly perturbed complex networks. *IET Control. Theory Appl.* **8**(16), 1639–1647 (2014)
9. Cai, C., Wang, Z., Xu, J., Liu, X., Alsaadi, F.E.: An integrated approach to global synchronization and state estimation for nonlinear singularly perturbed complex networks. *IEEE Trans. Cybern.* **45**(8), 1597–1609 (2015)
10. Cai, C., Xu, J., Liu, Y., Zou, Y.: Synchronization for linear singularly perturbed complex networks with coupling delays. *Int. J. Gen. Syst.* **44**(2), 240–253 (2015)
11. Peng, C., Li, F.: A survey on recent advances in event-triggered communication and control. *Inf. Sci.* **457–458**, 113–125 (2018)
12. Li, Q., Shen, B., Liang, J., Shu, H.: Event-triggered synchronization control for complex networks with uncertain inner coupling. *Int. J. Gen. Syst.* **44**(2), 212–225 (2015)
13. Liu, T., Cao, M., Persis, C.D., Hendrickx, J.M.: Distributed event-triggered control for synchronization of dynamical networks with estimators*. *IFAC Proc. Vol.* **46**(27), 116–121 (2013). (4th IFAC Workshop on Distributed Estimation and Control in Networked Systems)

14. Sivarajanji, K., Rakkiyappan, R., Cao, J., Alsaedi, A.: Synchronization of nonlinear singularly perturbed complex networks with uncertain inner coupling via event triggered control. *Appl. Math. Comput.* **311**, 283–299 (2017)
15. Schäcke, K.: On the kronecker product kathrin schäcke (2013)
16. Bhandari, M., Fulwani, D.M., Gupta, R.: Event-triggered composite control of a two time scale system. *IEEE Trans. Circuits Syst. II: Express Briefs* **65**(4), 471–475 (2018)
17. Abdelrahim, M., Postoyan, R., Daafouz, J.: Event-triggered control of nonlinear singularly perturbed systems based only on the slow dynamics. In: 9th IFAC Symposium on Nonlinear Control Systems. NOLCOS 2013, pp. 347–352. Toulouse, France (2013)

A New Compact Multilevel Inverter Design with Less Power Electronics Component



Siddharth Pachpor, Subhankar Dutta, M. Jagabar Sathik and Mudit Babel

Abstract Nowadays, multilevel inverter with reduced switch count is more attractive among the researchers due to the unsuitability of the conventional multilevel inverter where the installation space is constrained. Two different algorithms have been proposed for determining the magnitude of DC voltage sources. In this project, a compact version of multilevel inverter design is suggested with reduced and limited power electronic components, which are compared with conventional and recent multilevel inverter topology in terms of a number of levels, auxiliary diode, gate driver circuits, and blocking voltage of switches. Simulation software such as MATLAB/Simulink and laboratory workbench-based experimental test has been conducted. Finally, the comparison between the simulation output and experimental is discussed to prove the superiority of suggested topology.

Keywords Multilevel inverter · Blocking voltages · DC voltage sources · Switches · Gate driver circuits

1 Introduction

The Cascaded H-Bridge (CHB) topologies are appropriate choices for abnormal state applications from the perspective of measured quality and straightforwardness of control. The quantity of switching devices and other major components utilized by the converter plays a major role in determining the power quality. The reason being the power quality is directly proportional to the number of voltage levels [1]. Also, there is a significant increase in the cost if the number of components used is

S. Pachpor (✉) · S. Dutta · M. Jagabar Sathik

Department of Electrical and Electronics Engineering, SRM Institute of Science and Technology, Chennai 603203, India

e-mail: siddharthpachpor22@gmail.com

M. Babel

Department of Electronics and Telecommunication Engineering, Rajiv Gandhi College of Engineering and Research, Nagpur, India

increased. The control complexity also increases to a large extent and this decreases productivity of the converter [2]. The level of the H-bridge increases the THD value decrease by using lesser number of switching components and control circuits for these semiconductor switches. As a result of lesser number of switching devices and control circuit, the system is compact as well as it can be designed easily. Flying capacitor is used as clamping device in the FC multilevel converter [3]. Symmetric MLI Topologies with Reduced Switching Devices, a novel Multilevel Inverter Based on Switched DC Sources, are some of the major existing topologies [4].

2 The New Proposed Multilevel Inverter Topology

The suggested new design consists of two unidirectional power switches, one DC voltage source, three DC-link capacitors, two diodes, and an H-Bridge multilevel inverter as shown in Fig. 1. The capacitor and the diode are connected across each other. This presented combination is then connected in series with the main unidirectional switch. By suitable switching of U1 and B1 and S1–S4, this suggested fundamental structure can generate seven voltage levels of 0, $+V_1$, $-V_1$, $+(V_1 + V_2)$, $-(V_1 + V_2)$, $+(V_1 + V_2 + V_3)$, and $-(V_1 + V_2 + V_3)$ at the output. The following Table 1 demonstrates the diverse estimations of the output voltage levels in light of the distinctive states of switches [5] (Fig. 2).

As shown in Fig. 1, to Maximum voltage = $V_1 + V_2 + V_3$.

- In time interval (0, t_1), voltage level of 0 Volts is generated, one of the switches combination (S1, S4) or (S2, S3) is activated. (Suppose, switch S1 and S2 gets activated).
- Now, for (t_1, t_2), only switches U1, S1, and S4 are to be activated. This generates voltage V_1 at output and similarly, we use the switching logics to get different voltage outputs as shown in Table 1. Switches are actually nonideal and hence, a chance of occurrence of a phenomenon known as dead time exists in high probability at times $t_2, t_3 \dots t_8$ in a period of output. So, if switches S2 and S4 are activated, one of the voltages sources is thereby short-circuited by S2 and S4. These conditions happen at times $t_2, t_3 \dots t_8$ in a period of output. This

Fig. 1 New proposed fundamental unit of multilevel inverter

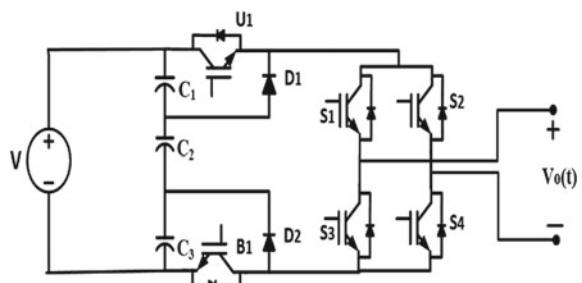


Table 1 Switching table for the new proposed cascaded MLI topology

State	U1	B1	S1	S2	S3	S4	V0
1	0	0	1	1	0	0	0 V
2	0	0	0	0	1	1	0 V
3	1	0	1	0	0	1	V_1
4	1	0	0	1	1	0	$-V_1$
5	0	1	1	0	0	1	$(V_1 + V_2)$
6	0	1	0	1	1	0	$-(V_1 + V_2)$
7	1	1	1	0	0	1	$(V_1 + V_2 + V_3)$
8	0	0	0	0	1	1	$-(V_1 + V_2 + V_3)$

1 = On state

0 = Off state

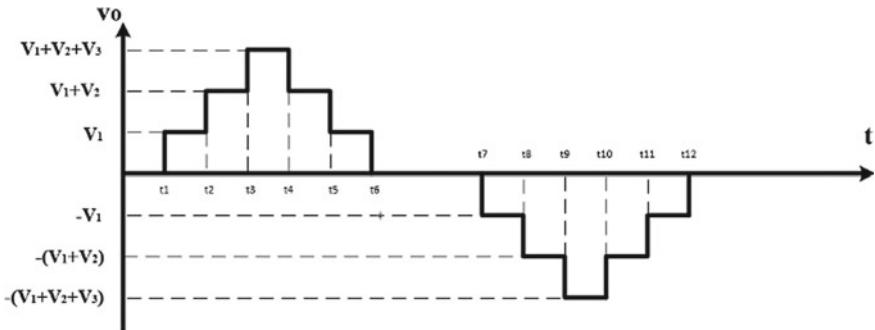


Fig. 2 Output voltage waveform for fundamental unit of the new proposed MLI

dead time occurs for a very small duration (in microseconds) [5]. This new proposed cascaded multilevel inverter formed using n fundamental basic units is presented in Fig. 3. The output voltages of different H-bridges are shown by $V_{o,1}$, $V_{o,2}$... and $V_{o,n}$

For this proposed new cascaded multilevel inverter, the Output voltage (V_o) is calculated as [5]

$$V_o = V_{o,1} + V_{o,2} + \dots + V_{o,n} \quad (1)$$

Equal voltage sharing of the capacitor voltages occurs and hence, this design is called as symmetrical multilevel inverter. For the proposed topology, the count of switches used (N_{switch}), number of levels (N_{level}), and the number of DC voltage sources (N_{source}), are written as equations [6]

$$N_{switch} = 6n, \quad (2)$$

$$N_{source} = n, \quad (3)$$

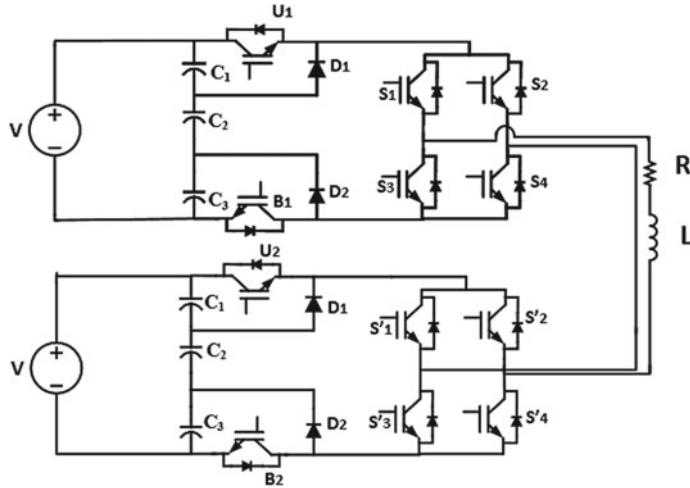


Fig. 3 Proposed new cascaded multilevel inverter

$$N_{\text{level}} = 6n + 1. \quad (4)$$

Similarly, for the proposed new topology, we used as an asymmetrical multilevel inverter, the following equations are written as equations:

$$N_{\text{switch}} = 6n, \quad (5)$$

$$N_{\text{source}} = n, \quad (6)$$

$$N_{\text{level}} = 7^n. \quad (7)$$

Multiple DC voltage sources for various increased levels of operation may be available in some systems such capacitors or batteries, etc. [6]. At the point when AC voltage is readily accessible, isolated transformers and rectifiers can be used in order to get DC voltage. These two algorithms demonstrate the different levels (even and odd), the magnitude of DC voltage sources or number of levels, and the number of switches [6]. Subsequently, graphs denote that new suggested design is the best topology given using 7-level MLI.

3 Proposed Algorithms

3.1 First New Proposed Algorithm

In the first algorithm, the magnitudes of the DC voltage sources are calculated by

$$V_{1,k} = V_{dc} \quad \text{For } k = 1, 2, 3, \dots, n \quad (8)$$

According to this algorithm,

$$N_{level} = 6n + 1 \quad (9)$$

The maximum amount of blocked voltage of the switches (V_{block}) and the variety of values for the number of DC voltage sources ($N_{variety}$) [7] is given by

$$V_{block} = 14n V_{dc} \quad (10)$$

$$N_{variety} = 1 \quad (11)$$

3.2 Second New Suggested Algorithm

Now, in the second algorithm, the magnitudes of the DC voltage sources are selected by

$$V_{1,k} = V_{2,k} = V_{2,k} V_{dc} \quad \text{For } k = 1, 2, 3, \dots, n \quad (12)$$

$$V_{block} = 14(2^{n-1})V_{dc}, \quad (13)$$

$$N_{variety} = 1, \quad (14)$$

$$V_{level} = 6(2^{n-1}) + 1V_{dc}. \quad (15)$$

4 Comparison Graphs Between Conventional and New Proposed Multilevel Inverter Topology

Finally, complete comparisons such as various parameters such as number of voltage levels, switches (IGBTs), and DC voltage sources are made. Similarly, graphs are also plotted. The comparison is between the proposed new MLI topology and the conventional topologies such as line R5 [8] and R2 [9] in Fig. 4, signifies that the

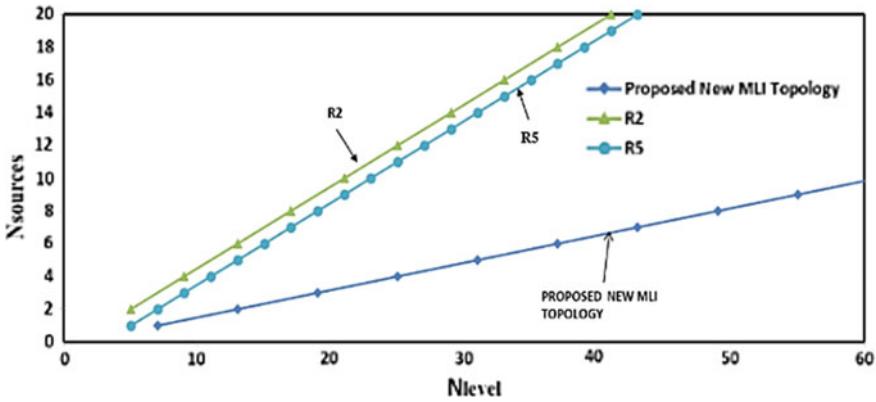


Fig. 4 $N_{sources}$ versus N_{level} comparison

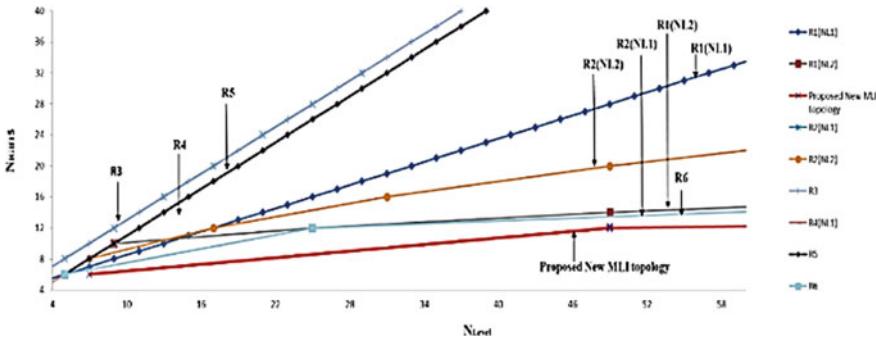


Fig. 5 N_{IGBTs} versus N_{level} comparison

new proposed MLI topology uses less number of sources and has more number of levels of the voltage output. Additionally, comparison between proposed New MLI Topology and the conventional topologies such as line R1(NL1) [2], R1(NL2) [2], R2(NL1) [9], R3 [10], R4 [11], R5 [8], and R6 [8] in Fig. 5 demonstrates that new Proposed MLI topology uses less number of switches (IGBTs). It also has more number of levels of the voltage output.

5 Simulation and Experimental Results

Fundamental frequency switching is one of the most common modulating techniques. In the presented structure, the fundamental frequency switching method (FFSM) has been utilized [12]. Also, calculation of switching angles for the elimination of the selected harmonics is not the primary aim of this proposed topology work. For an R-L load of 50 mH and $150\ \Omega$, we have chosen as a standard load for the testing

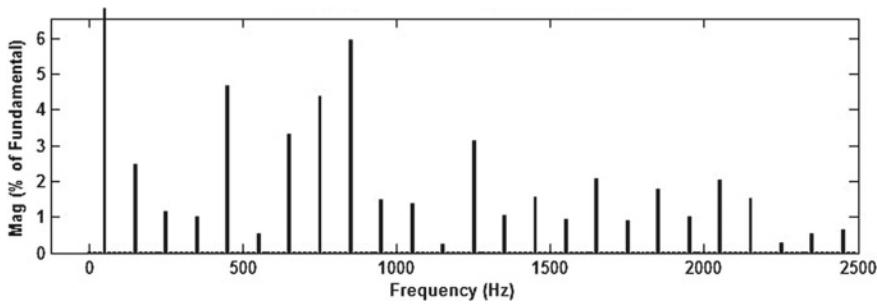


Fig. 6 THD waveform of the new suggested topology

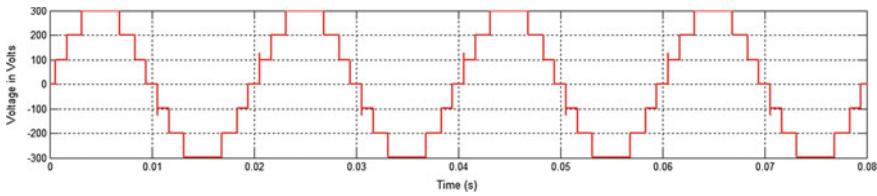


Fig. 7 Simulated graph for V_{dc} versus time results for basic unit of 7-level cascaded inverter based on proposed topology

purpose. We have also used 25 V, 4700 microfarad DC-link capacitors. IGBT is used because of its high voltage capability [12]. To investigate of the recommended MLI, the simulation and experimental results are performed for a 7-level multilevel inverter on MATLAB/Simulink software program.

Estimated output voltage frequency = 50 Hz.

Simulation software used is MATLAB/Simulink programming,

$$THD = \frac{\sqrt{(\sum_{H=1,3,5}^{\infty} V_{oh})}}{V_{o1}} = \sqrt{\left(\frac{V_{orms}}{V_{o1}}\right)^2 - 1} \quad (16)$$

In this above relation, h denotes corresponding harmonic order. Fundamental frequency is denoted by subindex 1. Hence, V_{oh} is the root mean square (RMS) of the n th-order harmonic. V_{o1} is the fundamental voltage of the output voltage waveform [5]. This inverter consists of one basic proposed unit, and can be cascaded into n units. Suppose we chose value of voltage source as 3 V in the first unit into our cascaded multilevel inverter design, 21 V will be the value for the second voltage source unit. This inverter can generate 343 levels with the maximum value of 172 V at the output. MOSFETs used on the prototype are IRF500 with two internal diodes and the following graph output is generated. Using XC3S100E FPGA, all switching patterns has been generated. THD or the distortion factor analysis graph is shown in Fig. 6 (Figs. 7 and 8).

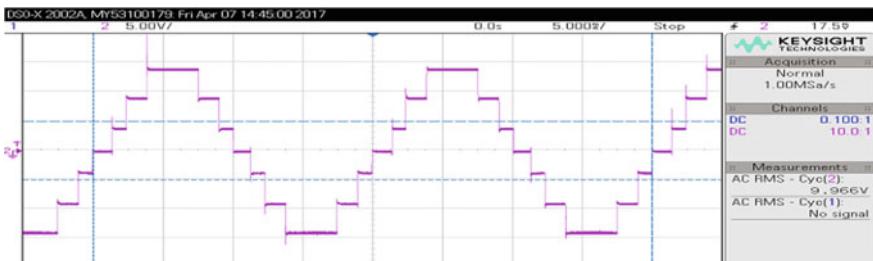


Fig. 8 Experimental results for V_{dc} versus time for basic unit of 7-level cascaded inverter based on proposed topology

6 Conclusions

The prospect of the proposed topology can also be used for high voltage applications purposes [13]. Through the results via the MATLAB simulation and hardware implementation, it is observed that staircase waveform quality waveform is obtained. The topology is also helpful to reduce the dv/dt stresses as the number of switch count has reduced significantly as compared to conventional topologies. There is also a significant increase in number of levels at the voltage output in the proposed topology as seen from the devised algorithms.

Also, the EMC (electromagnetic compatibility) issues can be reduced using the proposed topology. Comparison graphs are constructed with respect to some sample conventional topology to get an idea of number of voltage levels, diode, blocking voltage of switches and subsequently, our suggested topology is considered the best. The major applications of the proposed topology can be in Electric Hybrid vehicles and Renewable energy system.

Acknowledgements We also take the opportunity to express a deep sense of gratitude and deep regards to our program guide Dr. Arun Noyal Doss, M.E. Ph.D and Dr Prabodh Pachpor, Ph.D for their able guidance and support. They not only motivated throughout the project but also additionally provided with constructive criticism.

References

1. Meynard, T.A., Foch, H.: Multi-level conversion: high voltage choppers and voltage-source inverters. In: 23rd Annual IEEE Power Electronics Specialists Conference, pp. 397–403 (1992)
2. Samanbakhsh, R., Taheri, A., Bahmani, H., Keshavarz, A.: Reduction of power electronic components in multilevel converters using new switched capacitor-diode structure. *IEEE Trans. Ind. Electron.* 7204–7214 (2016). <https://doi.org/10.1109/tie.2016.2569059>
3. Babei, E., Dehghan, A., Sabahi, M.: A new Topology for multilevel inverter considering its optimal structures. *Electr. Power Syst. Res.* (2013)

4. Wang, T., Zhu, Y.: Analysis and comparison of multicarrier PWM schemes applied in H-bridge cascaded multi-level inverters. In: 5th IEEE Conference on Industrial Electronics and Applications, pp. 1379–1383 (2010)
5. Babaei, E., Laali, S., Bahravar, S.: A new cascaded multi-level inverter topology with reduced number of components and charge balance control methods capabilities. *Electr. Power Compon. Syst.* (2015)
6. Babaei, E., Laali, S., Bayat, Z.: A single-phase cascaded multilevel inverter based on a new basic unit with reduced number of power switches. *IEEE Trans. Ind. Electron.* (2015)
7. Babaei, E., Laali, S.: Optimum structures of proposed new cascaded multilevel inverter with reduced number of components. *IEEE Trans. Ind. Electron.* (2015)
8. Alishah, R.S., Nazarpour, D., Hosseini, S.H., Sabahi, M.: Reduction of power electronic elements in multilevel converters using a new cascade structure. *IEEE Trans. Ind. Electron.* **62**(1), 256–269 (2015)
9. Busquets-Monge, S., Caballero Diaz, L.: Switching-cell arrays—an alternative design approach in power conversion. *IEEE Trans. Ind. Electron.* (2018)
10. Ebrahimi, J., Babaei, E., Gharehpetian, G.B.: A new multilevel converter topology with reduced number of power electronic components. *IEEE Trans. Ind. Electron.* **59**(2), 655–667 (2012). <https://doi.org/10.1109/tie.2011.2151813>
11. Ebrahimi, J., Babaei, E., Gharehpetian, G.B.: A new topology of cascaded multilevel converters with reduced number of components for high-voltage applications. *IEEE Trans. Power Electron.* **26**(11), 3109–3118 (2011). <https://doi.org/10.1109/tpel.2011.2148177>
12. Yang, Y., Ma, K., Wang, H., Blaabjerg, F.: Instantaneous thermal modeling of the DC-link capacitor in photovoltaic systems. In: 2015 IEEE Applied Power Electronics Conference and Exposition (APEC), pp. 2733–2739 (2015)
13. Kotsopoulos, A., Duarte, J.L., Hendrix, M.A.M.: A predictive control scheme for dc voltage and ac current in grid-connected photovoltaic inverters with minimum dc link capacitance. In: Proceedings of 27th Annual Conference IEEE Industrial Electronics Society, vol. 3, pp. 1994–1999 (2001)

Detection of Emergency Vehicles Using Modified Yolo Algorithm



Somya Goel, Abhishek Baghel, Aprajita Srivastava, Aayushi Tyagi and Preeti Nagrath

Abstract This paper is a brief overview of the topic of object recognition and how it can be used to detect different objects present on a busy road. Using this technology, we aim to detect emergency vehicles present on the road so that a free path can be created for them, which would be able to save precious time in time-sensitive situations. The introduction to object recognition includes what is the work done and how it can be executed in the best way possible in an error-free manner. The best results are depicted in our research methodology. In this paper, we compared different methodologies available for object recognition such as YOLO algorithm, Region-based convolutional networks (R-CNN), and Single-Shot Detector (SSD) and ultimately chose YOLO algorithm, which from our results would work optimally in a situation where emergency vehicles need to be detected on a busy road.

Keywords Object recognition · YOLO · R-CNN · Emergency vehicle detection · Deep learning · SSD

1 Introduction

Object recognition is one of the most exciting areas in AI and computer vision. The ability to identify all the objects in a background seems to be no more a secret of evolution. Object recognition is a technique used to identify things or entities. It is widely used in the area of artificial intelligence in which this technique is used by the robots to identify various objects around it and react accordingly. Sometimes, object recognition can be confused with objection detection. Object detection is the process in which the object is founded in some image or 2D object.

Emergency vehicles such as fire trucks and ambulances use sirens, lights, and different color combinations to alert nearby traffic, where they are in a need of moving quickly. In spite of this, lots of drivers are still unaware of these vehicles.

S. Goel · A. Baghel · A. Srivastava (✉) · A. Tyagi · P. Nagrath
Computer Science Department, Bharati Vidyapeeth's College of Engineering, New Delhi, India
e-mail: aprajita4@gmail.com

This can be due to a variety of reasons including inattention or even ignorance. Alerting other drivers and pedestrians is very important to reduce accidents on the road and creating a path for these emergency vehicles. Also, for self-driving cars, it is essential to distinguish these vehicles from normal traffic to give them space and clear passage.

By using the technology of object recognition, we can use a simple camera system to detect these vehicles present on the road. These vehicles would include ambulances, police vehicles, and even fire trucks that need to work in a time-sensitive manner. An emergency vehicle detection system has various applications, but this project focuses on the subject of awareness of emergency vehicles moving in traffic to human or nonhuman drivers.

In this paper, we'll compare different algorithms and models available for the purpose of object recognition. These models involve deep learning and will need to be trained for any particular dataset.

1.1 Complexities of Object Detection for Emergency Vehicle Detection

Humans can recognize objects so effortlessly and instantly even if picture of the objects varies in different viewpoints. Objects can even be identified when they are partially blocked from view but if we actually realize object identification is a very complex process especially for implementation of algorithmic description on machines. Some of the complexities are described below:

- ***Image Model Spaces***:—Sometimes the 3D images appear as 2D images. If models are 3D and projection effects of different dimensions cannot be avoided, then the problem becomes more complicated.
- ***The number of the classes in the database***:—If the database from which the object identification is taking place is huge and contains a large number of object classes, then the time taken by the machine to identify those objects will increase. (This problem will arise if the database is used for object recognition).
- ***The possibility of occlusion***:—Occlusion means blocking. It happens when certain features are missing from the objects, that are important for identification, and some unexpected features are present instead. This creates many problems in identification.
- ***Rotation or scale***:—The objects may be present with different orientations or of different sizes. If these orientations and scales are not accurately mentioned in the database, then it can cause misinterpretation of no interpretation at all of the objects.

1.2 *Introduction to Object Recognition*

Object recognition is the method of identifying the object, i.e., what exactly it is as well as finding the location of the object. Object recognition can be achieved by using various methods/algorithms of machine learning and deep learning. An object recognition uses a new class of features of the local image. These features do not change with translation, image scaling, and rotation [1]. There are various techniques to do object recognition but the basic steps involved remain the same.

The basic steps involved in the identification of objects are [2] the following:

- **Object Representation**—It tells how objects are represented in the database. The important features or attributes of objects that must be present in these models? The representation of all the necessary details of the object those are required for recognizing it.
- **Feature detection and Matching**—The features should be detected and then extracted and compared with details mentioned in the database.
- **Hypotheses formation**—The hypothesis formation is a heuristic to decrease the size of the search space.
- **Object Verification**—The last step of object recognition. The presence of each object can be identified by using their models. a plausible hypothesis must be examined to verify the presence of the object.

Object recognition can be done in various ways such as by using MATLAB, 3D detection method, neural networks, and various other algorithms. But whatever algorithm is used in object recognition, they all need to have some basic characteristics such as [3]

- **Reliability**: The algorithm must be reliable, it should be able to capture the mapping function that is robust to image variations and noise. It should perform well on a set of having a different distribution.
- **Speed**: The applications may be made to work in live environment, this implies that speed of the image identification process should be quick so that results can be shown in real time.
- **Automation**: It should work smoothly without any additional user interaction.

Object recognition has remained an important field of research to find out the efficient methods of performing it. Nowadays, AI and further robotics are developing at a faster rate. Robotics is having a wide range of applications and object recognition is the most important part of the robotics without which it is difficult to expand robotics. In the past few years, object recognition has become a vital component in various applications [4] such as

- Biometric scanning.
- Robotics.
- Human–machine interaction.
- Industrial machine inspection.
- Pharmaceutical analysis.

- Content-based image retrieval (CBIR).
- Automatic driving cars.

1.3 *Organization of the Paper*

Section 1 of the paper deals with the introduction of the research done and explains the problem statement in a detailed manner. Section 2 throws light on the topic of object recognition, which is the foundation of our model. Section 4 tells the similar work that has been done in the past, which were used as a reference for our paper. Section 5 deals with the research methodology and Sect. 6 explain the different algorithms available for object recognition. Section 7 of the paper compares these different algorithms and eventually draws a result. Section 7 also concludes the paper and also adds the future scope for this research work.

2 Literature Survey

Lowe et al. [1] talks about the various techniques that are used in object recognition. The best one is object localization, it is particularly helpful in local image features. Some experimental results were observed and could be applied really well on the object recognition models. Kurian et al. in 2011 [3] told about how object recognition is used in video sequences and audio detection at a good efficiency. It throws light on the two wide models, i.e., Appearance-based and Model-based. It talks about various other models like Contour based, template based etc. Markus et al. in 2013 [5] explained the application in talks, traffic sign recognition. This is one application in which object recognition is both applicable and important. It showed that the current models could reach an efficiency of about 95%. It also highlights the parameters on which the traffic dataset was programmed. Ziegler et al. [6] wrote about a historical journey with the Mercedes S-Class and various ways of detection, like radar sensors, digital maps, etc., which can prove to be good indicators of automated driving.

Ros et al. in 2015 [7] gave insight into autonomous driving and the technique of mapping to perform the same. The author performed the experiment on the KITTI dataset and got the best results on various pointers like timing information. As per Uijilings et al. in 2015, [8] highlights a selective search that can combine the best of both worlds by coalescing both exhaustive search and segmentation thereby leading to better results.

Redmon et al. in 2016 [9] and this group of researchers finally proposed the YOLO model (You Only Look Once) and gave both its advantages and disadvantages. It divides objects into bounding boxes and then applies them to real-time situations to observe the best results in terms of accuracy. It combined the likes of R-CNN with the YOLO algorithm. As per the information provided by the U.S. Department of Transportation, more than 30,000 people lost their lives in car accidents in the U.S.

alone in 2009 [10]. Autonomous driving holds the power to significantly bring down that number. Autonomous driving could also help people from wasting their precious time in commuting which could increase their productivity. Also, self-driving cars are a more sustainable option for the environment in the long run. The technology of 3D object detection can be used in autonomous driving. Urmson [11] in his paper throws light on the project of self-driving cars that was undertaken by Google. His paper stated that detailed and high-quality object recognition is not required, so far, for the purpose of creating self-driving cars. Teichman et al. [12] in their paper, state that by using a laser-based object recognition system, we would not require an object class model. The algorithm described here can be broken down into three parts: Segmentation, tracking, and classification. Classification of objects in a video is done by a boosting algorithm applied across several high dimensional descriptor spaces, which encode size, shape, and motion properties. Spinello et al. [13] tell us through an example that method based on laser uses a model-free segmentation technique and tracking, whereas the tracking-by-detection method involves passing individual frames through the detector, then tracking the resulting detections from the individual results.

3 Research Methodology

Object recognition is a part of machine learning and artificial intelligence, nowadays, which is a fast-growing field of computer science. It is a technique used to identify objects in images or videos, along with some real-time applications. The human beings identify a variety of things in images instantly without taking much effort, in spite of the difference in image sizes or even images are rotated or translated. This is a challenge in this domain of computer vision.

The various applications [14] of object recognition are the following:

1. Face detection.
2. Pedestrian detection.
3. Automated vehicle parking systems.
4. Robot localization.
5. Automatic target recognition.

As shown in Fig. 1, the different objects are classified and given various colors and the distances from each are measured with respect to the object in consideration.

In recent times autonomous driving has been a major focus of research and development. There are a number of methods to work on this path. There are various areas to keep in mind like behaviors, patterns, street signs, parking facilities, etc. [4].

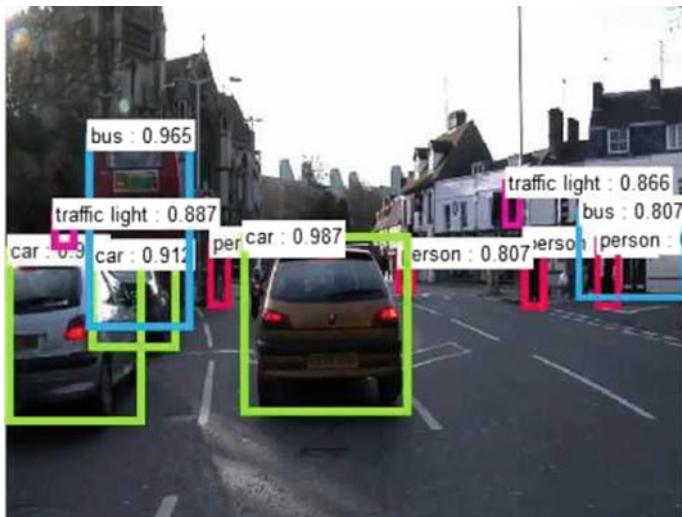


Fig. 1 An example of autonomous driving (Source <https://www.xsens.com/customer-cases/vector-ai-shaping-the-future-for-autonomous-driving/>)

3.1 *There Are Basically Two Wide Methods for Recognition*

Appearance-based recognition: In this method, we use images to gather data about the difference on the basis of appearances like changing in lighting or color, in viewing direction, and varied sizes and shapes. It makes use of different algorithms and steps that include edge matching to produce excellent measurements. Another technique is divide-and-conquer search wherein the cells are divided into sub-cells and they work on each sub-cell recursively, hence it is assured to find all matches of a condition. Gray scale matching and gradient matching are also these algorithms.

Feature-based recognition: It is a search method that finds feasible matches between an object and its features and other image features. The major drawback is that the position of the object is to be kept steady. The various forms are interpretation trees that are searching through a tree, hypothesize and test, clustering, and geometric hashing.

There are two broad processes currently in use, which are the following:

1. **Supervised Method:** It is a laser-based algorithm based on three main components: segmentation, tracking of vehicles and objects and lastly, classification of the track. Objects are segmented using information obtained about their depth and thereafter tracked with a filter. Classifying the tracks is done using a boosting algorithm [5].
2. **Semi-supervised Method:** It is a tracking-based semi-supervised learning. It will make use of a classifier to collect new and useful instances by collecting the tracking information from given datasets. It is also a model that undergoes object

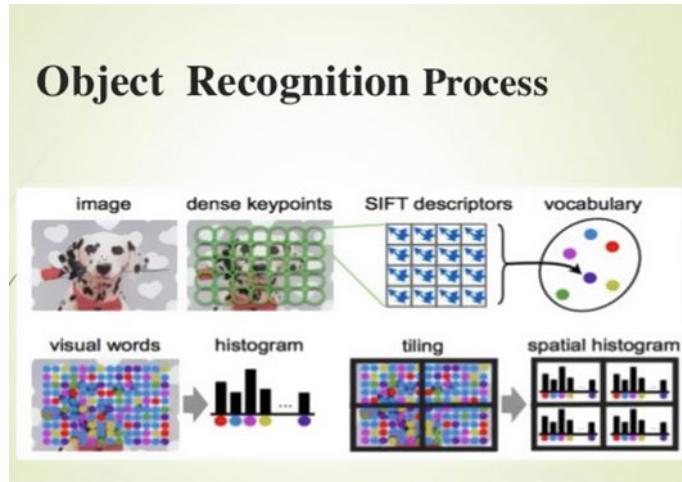


Fig. 2 Object recognition process (Source <https://www.slideshare.net/sanicorreya/object-recognition-57543686>)

detection. We set an object plane in order to set objects close to it. Our image is processed and ground plane is said to be orthogonal to an image plane and some distance down. The steps are generating 3D object proposals with respect to semantic segmentation, shape, instance segmentation, context, and location. 3D proposal learning and inference, CNN scoring of top proposals and other implementation details [15].

The classification types used are boosting frameworks used to classify individual segments. The boosting framework has two classifiers [16]. The first one is a segment classifier which details about the appearance of the given object at a moment. There is a holistic classifier, which talks about various parameters like—speed, acceleration, and various other properties (Fig. 2).

4 Models

4.1 R-CNN—Region-Based Convolutional Network

In R-CNN, a selective method is applied for searching. It was developed by Uijlings et al. (2013) so that it could perform object recognition in an image to find and capture the correct location of an object with greater speed. It is a tweak to sliding-window algorithms. The algorithm uses small, special regions in a picture generated after preprocessing, merges them in hierarchical group and then performs classification on those groups. The results show that the final group is a box containing the full



Fig. 3 Visualization of the region proposals of the algorithm (*Source* Selective Search application, top: visualization of the segmentation results of the algorithm, down: visualization of the region proposals of the algorithm. *Source* Uijlings et al. (2013))

picture. The regions which are detected by R-CNN are coalesced according to various similarity metrics and color spaces. The hierarchical group obtained is a smaller number of regions, which may contain a final object by merging small regions as shown in Fig. 3. The selective search method is combined with R-CNN model (Girshick et al. 2016) to detect deep learning and region proposals to detect the object in these regions. The architecture of a typical R-CNN is shown in Fig. 4.

The best R-CNN models have achieved a 31.4% mAP score over the 2013 ImageNet dataset and 62.4% mAP score over the PASCAL VOC 2012 test dataset. R-CNN has further improved versions such as Faster-CNN. Faster R-CNN is broadly classified into two kinds of networks: Region Proposal Network (RPN), which is used to generate probable regions and a consequent network using these regions to recognize objects. The major difference between RPN and Fast R-CNN is that R-CNN makes use of selective search to form region proposals.

4.2 You Only Look Once (Yolo)

The YOLO model [9] is a direct predictor that does so with the help of bounding boxes, and also class probabilities are used with a single network to perform a single and efficient evaluation (Fig. 5). The YOLO model shows real-time evaluation unlike

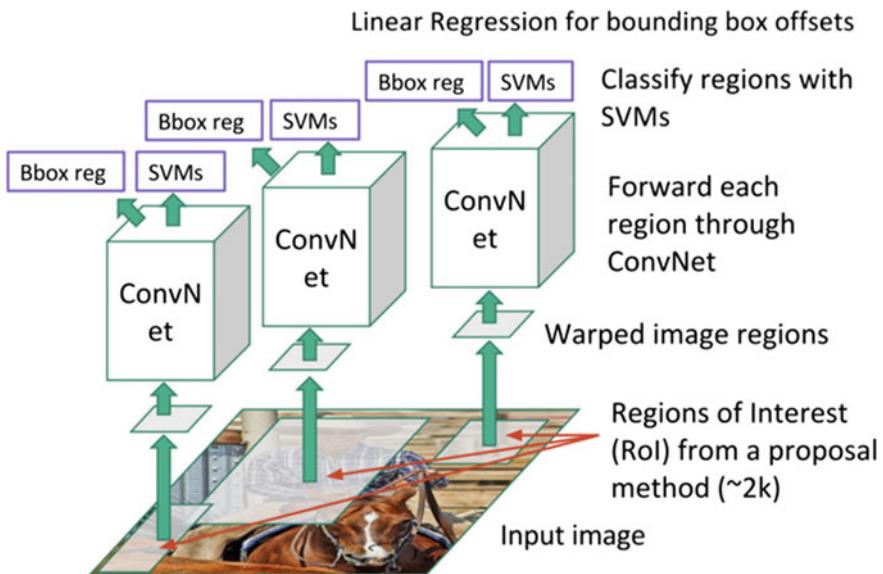


Fig. 4 Region-based Convolutional Network (R-CNN). Each probable region expected to contain the object is passed to the network instead of the possible regions of the picture (*Source* <https://towardsdatascience.com/deep-learning-for-object-detection-a-comprehensive-review-73930816d8d9>)

many other algorithms like R-CNN and SSD. First, the model will take an input image. It is further divided into an $S \times S$ grid. A confidence score is given to each bounding box labeled as B by the grid. This confidence measures the probability of the object belonging to a specific class that is successfully detected object belonging to a specific class as shown in Fig. 6. The count of convolution layers is 24 and the fully connected layers are 2. Reduction in layers with filters of 1×1 filters⁴, and then division of convolutional layers in grids of 3×3 layers replace the model that is initially proposed. There is also another YOLO model, the Fast-YOLO model. This is a comparatively lighter version containing very few filters and almost nine convolutional layers. In the ImageNet dataset, most of the convolutional layers are then trained ahead using classification. There is an addition to the previous network model with four convolutional layers and two fully connected layers. A typical YOLO architecture is shown in Fig. 6.

The YOLO model has a 63.7% Mean average precision (mAP) when analyzed over the 2007 Pa VOC dataset and a 57.9% Mean average precision (mAP) results are compared for the 2012 Pa VOC dataset. There are other versions of YOLO that give comparatively lower accuracy score but have a faster speed for real-time applications. Improved versions of YOLO include YOLOv2, Fast-YOLO, etc. Fast-YOLO processes about 140 frames per second (even faster than real time) of an image

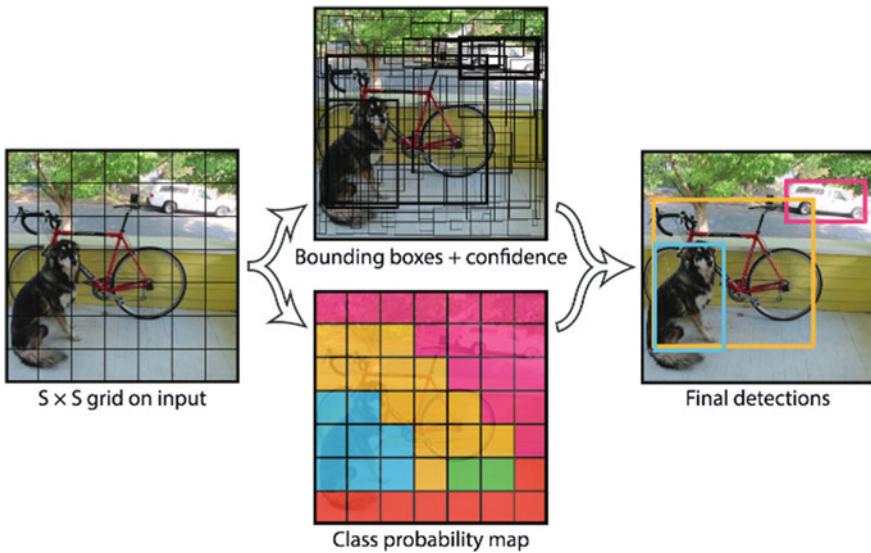


Fig. 5 Working of YOLO algorithm. The input image is divided into an $S \times S$ grid, B bounding boxes are predicted for each cell and a class is predicted for each box among C classes. Only the most confident boxes are displayed (Source <https://stackoverflow.com/questions/33947823/what-is-semantic-segmentation-compared-to-segmentation-and-scene-labeling>)

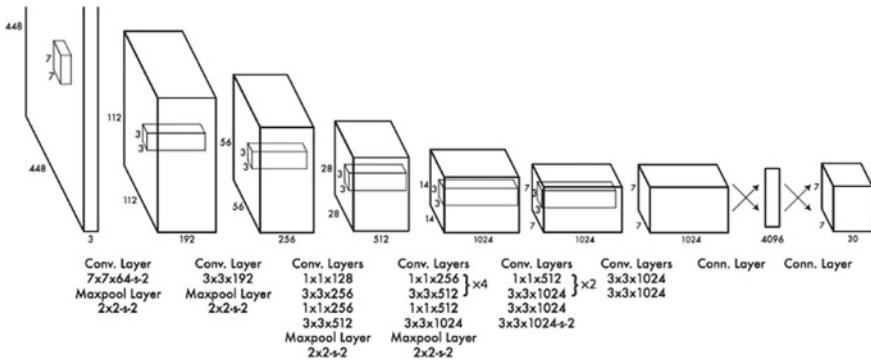


Fig. 6 YOLO architecture is made of 24 convolutional layers one after the other and 2 fully connected layers in the end (Source <https://towardsdatascience.com/yolo-you-only-look-once-real-time-object-detection-explained-492dc9230006>)

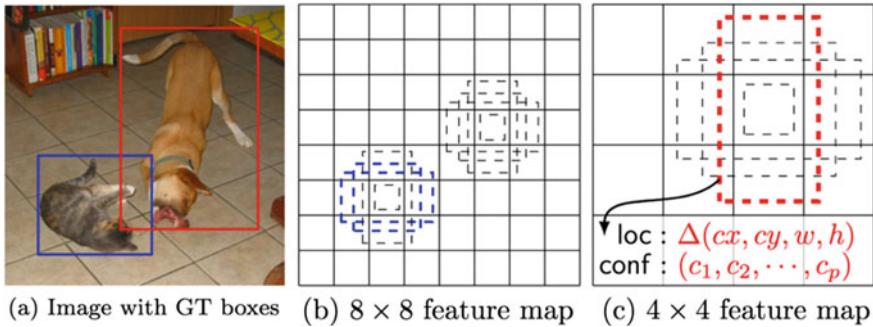


Fig. 7 Working of SSD Framework. **a** The model takes as input an image and its ground truth bounding boxes. Boxes having different aspect ratios are fixed by the different features. While training the boxes are modified to best match the ground truth (*Source <https://towardsdatascience.com/understanding-ssd-multibox-real-time-object-detection-in-deep-learning-495ef744fab>*)

as compared to about 50–60 frames per second processed by YOLO algorithm. The accuracy is compromised for speed in the Fast-YOLO algorithm.

4.3 Single-Shot Detector (SSD)

Similar to the YOLO model, Single-Shot Detector (SSD) was developed by Liu et al. [24] developed a model to predict and define all the bounding boxes at once and the corresponding class probabilities with the help of an exclusive CNN architecture. The model will take a particular image (10×10 , 5×5 and 3×3) as input, which passes through a number of convolutional layers having different sizes and number of the filters as per usage. Feature maps that are formed from convolutional layers at varied positions of the network are used for correct prediction of the bounding boxes. They are processed using a specific convolutional layer which consists of 3×3 filters also known as the extra feature layers to predict a set of bounding boxes as shown in Fig. 7.

The SSD models are trained with the 2015 COCO dataset and 2007, 2012 Pa VOC datasets with augmentation of data, mAP score—82.2% (2012 Pa VOC test dataset); mAP score—83.2% (2007 Pa VOC test dataset).

5 Analysis Various Models and the Result

There are a lot of object detectors that are present and the parameters to compare are many. Each parameter gives a new dimension to object recognition. However, speed and accuracy play the most important roles in object detection. It is tough to

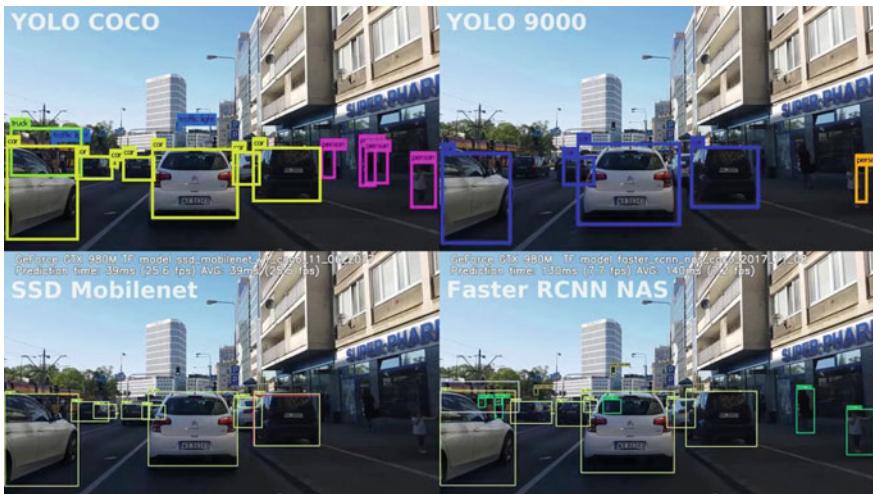


Fig. 8 Object recognition using different algorithms (Source <https://www.youtube.com/watch?v=ZKUUD1iXLQ0>)

say which detector is the best, instead, we must look at which parameters are the fundamentals in the success of a detector (Fig. 8).

Accuracy Versus Frames Per Second (FPS)

In the case of the various object recognition algorithms, the parameters measured are as follows:

1. Accuracy.
2. Frames per second.
3. IoU (Intersection over Union).
4. Real-time analysis.

Accuracy: An object detector's accuracy is measured by the IOU criteria. The measure normalizes the IoU with the use of boxes for segmenting data to generate a measure for accuracy (range—0–1) and gives a precise measurement.

Frames per second: The frames per second specify how many frames can be captured in the bounding boxes used in different object detection algorithms. This is a good choice of parameter for determining the speed of the algorithm.

IoU: This is a technique used to measure accuracy in object detectors. Mostly it is to be used when we use bounding boxes in our algorithm.

Real-time Analysis: Real-time analysis is a way to process the data in real time because we process applications in real-time.

In the following graphs, we compare all the following algorithms based on accuracy and frames per second as these two are the major factors in deciding an algorithm (Figs. 9 and 10).

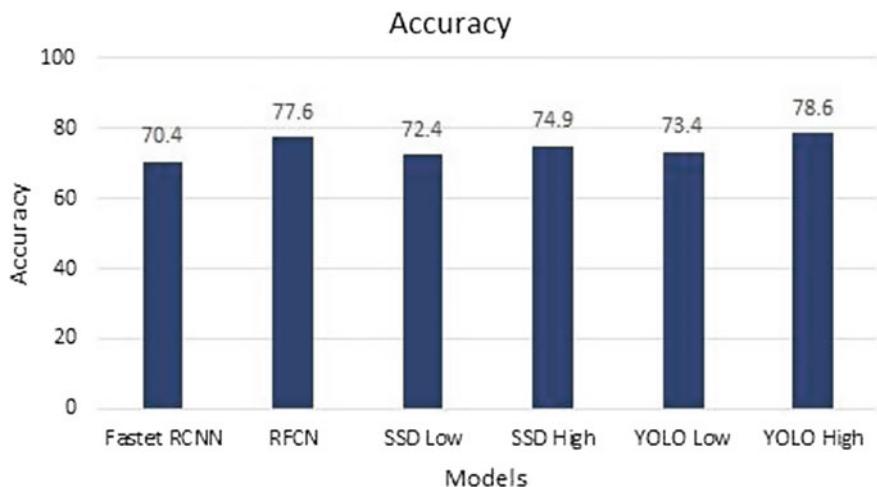


Fig. 9 Accuracy comparison for different object recognition algorithms (*Source* https://medium.com/@jonathan_hui/object-detection-speed-and-accuracy-comparison-faster-r-cnn-r-fcn-ssd-and-yolo-5425656ae359)

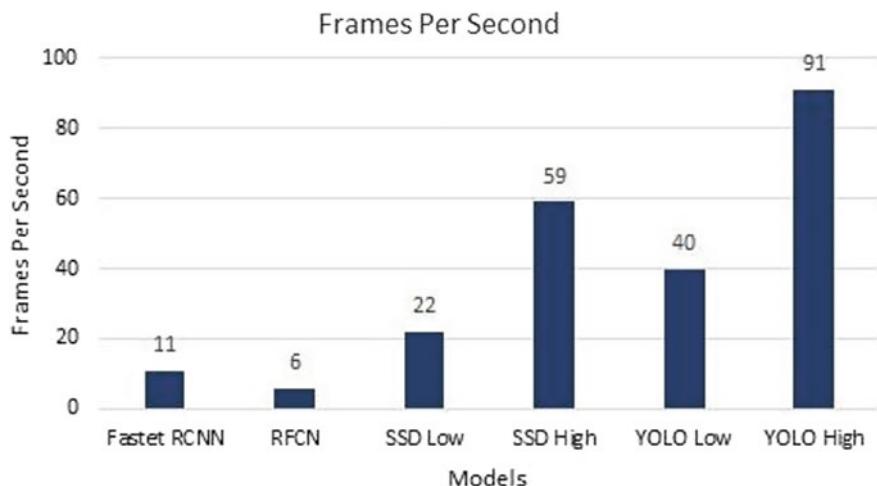


Fig. 10 Comparison of frames processed per second for different object recognition algorithms (*Source* https://medium.com/@jonathan_hui/object-detection-speed-and-accuracy-comparison-faster-r-cnn-r-fcn-ssd-and-yolo-5425656ae359)

Result:

The above two graphs show that YOLO algorithm is the most optimal choice for this particular problem statement due to the following reasons:

1. YOLO algorithm has very high values of mAP (Mean average precision), i.e., mAP = 63.4%, whereas the other algorithms except R-CNN do not show such good results.
2. The frames per second in case of the YOLO algorithm are lesser, i.e., FPS = 45. Other algorithms such as Fast-YOLO have a better frame rate, i.e., FPS = 155.
3. YOLO and Fast-YOLO are the only algorithms that work on real-time applications, which is imperative for this particular problem statement.

6 Proposed Model

The model is a 2-step ML pipeline.

In the first phase, the Fast-YOLO algorithm is applied to detect the position of various vehicles in the video as described earlier. This step outputs a $7*7*30$ tensor value containing the position of the objects relative to the image along with the confidence score for each object class detected.

This value is passed to an interface which crops the image of the vehicle detected in the previous step from the tensor received. It also includes a preprocessing step on the image to change the dimensions of the image into a fixed size before passing it to the next phase.

The second phase of the pipeline is a pretrained CNN that is a simple vehicle classifier that takes input an image of a vehicle and outputs a one-shot vector identifying the class of vehicle. The model classifies a vehicle into four classes, namely Police, Fire truck, Ambulance, and Others. Its architecture is as follows:

- The first layer is a convolution layer having $32 \ 3 \times 3$ filters and activation function as relu.
- The second layer is a nonlinear downsampling MaxPooling layer.
- The third layer is a convolution layer having $64 \ 3 \times 3$ filters and activation function as relu.
- The fourth layer is a nonlinear downsampling MaxPooling layer.
- The fifth layer is a convolution layer having $64 \ 3 \times 3$ filters and activation function as relu.
- The sixth layer is a fully connected layer having 150 nodes with sigmoid activation function.
- Dropout of 0.5 is applied to prevent overfitting

The output layer consists of four nodes, each for one classification class. It is a densely connected layer having a “softmax activation function”.

A major advantage of this pipeline is that In case other classes of vehicles are to be added we can simply re-train the second phase of the pipeline.

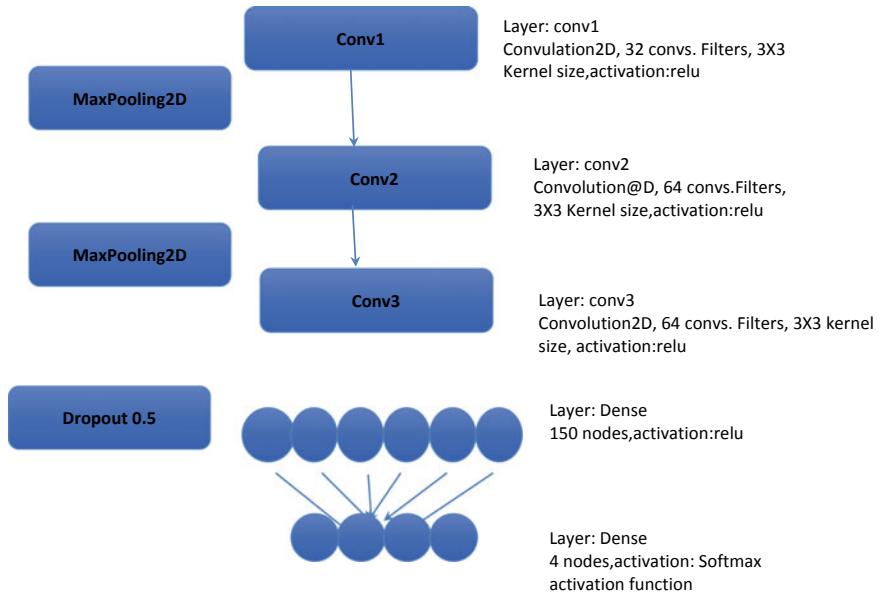


Fig. 11 The architecture of second-phase vehicle classifier

This proposed architecture can be further understood by the following figure [17] (Fig. 11).

7 Conclusions and Future Scope

The conclusion of the study or comparison is that even though YOLO has an excellent precision rate, the value for frames per second hampers the speed of the algorithm and doesn't let the algorithm work to its best utility. A good model would have the precision of YOLO and the frame rate of Fast-YOLO.

The intermediate between the two phases of the pipeline can be used to crop out the images of various objects detected by the YOLO algorithm and preprocess the cropped image to a dimension suitable for the respective classifier and pass them to the respective classifiers of object classes to a more detailed classification.

Detailed classifiers for various objects are available and can be used.

Since all the classifiers in the second phase work independent of each other. They can be easily scaled to multicore processors for optimum performance.

In the project ahead, the plan is to find the best ways to segment, track and classify various objects basically in the field of autonomous driving. Object recognition is possible with the OpenCV library of Python and we intend to utilize the same for our project's benefit, the various algorithms and models for object recognition will

enable us to present a working and efficient prototype of our own, which would be susceptible to errors, will be comparatively faster, and also be a good determiner for the field of autonomous car driving, which is prevalent nowadays, and it would help to prevent accidents and mishappenings in the future.

References

1. Lowe, D.G.: Object recognition from local scale-invariant features. In: The Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999, vol. 2, pp. 1150–1157. IEEE (1999)
2. http://www.cse.usf.edu/~r1k/MachineVisionBook/MachineVision.files/MachineVision_Chapter15.pdf
3. Kurian, M.Z.: Various object recognition techniques for computer vision. *J. Anal. Comput.* **7**(1), 39–47 (2011)
4. Bjorn, V.: One finger at a time: best practices for biometric security. Banking Information Source (Document ID: 1697301411) (2009)
5. Mathias, M., Timofte, R., Benenson, R., Van Gool, L.: Traffic sign recognition—how far are we from the solution? In: The 2013 International Joint Conference on Neural Networks (IJCNN), pp. 1–8. IEEE (2013)
6. Ziegler, J., Bender, P., Schreiber, M., Lategahn, H., Strauss, T., Stiller, C., Dang, T., Franke, U., Appenrodt, N., Keller, C.G., Kaus, E.: Making Bertha drive—an autonomous journey on a historic route. *IEEE Intell. Transport. Syst. Mag.* **6**(2), 8–20 (2014)
7. Ros, G., Ramos, S., Granados, M., Bakhtiari, A., Vazquez, D., Lopez, A.M.: Vision-based offline-online perception paradigm for autonomous driving. In: 2015 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 231–238. IEEE (2015)
8. Uijlings, J.R., Van De Sande, K.E., Gevers, T., Smeulders, A.W.: Selective search for object recognition. *Int. J. Comput. Vision* **104**(2), 154–171 (2013)
9. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788 (2016)
10. U.S. Department of Transportation. Traffic safety facts 2009 (early edition)
11. Urmson, C.: The Google self-driving car project. In: Talk at Robotics: Science and Systems (2011)
12. Teichman, A., Levinson, J., Thrun, S.: Towards 3D object recognition via classification of arbitrary object tracks. In: International Conference on Robotics and Automation (2011)
13. Spinello, L., Luber, M., Arras, K.O.: Tracking people in 3D using a bottom-up top-down detector. In: International Conference on Robotics and Automation (2011)
14. <https://www.mathworks.com/solutions/deep-learning/object-recognition.html>
15. Chen, X., Ma, H., Wan, J., Li, B., Xia, T.: Multi-view 3d object detection network for autonomous driving. In: IEEE CVPR, vol. 1, no. 2, p. 3 (2017)
16. Chen, X., Kundu, K., Zhang, Z., Ma, H., Fidler, S., Urtasun, R.: Monocular 3d object detection for autonomous driving. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2147–2156 (2016)
17. Liu, D., Wang, Y.: Monza: image classification of vehicle make and model using convolutional neural networks and transfer learning (2017)
18. Khurana, K., Awasthi, R.: Techniques for object recognition in images and multi-object detection. *Int. J. Adv. Res. Comput. Eng. Technol. (IJARCET)* **2**(4), 1383 (2013)
19. The PASCAL Visual Object Classes Homepage Host.robots.ox.ac.uk. <http://host.robots.ox.ac.uk/pascal/VOC/> (2018)
20. ImageNet. Image-net.org. <http://www.image-net.org/> (2018)

21. COCO—Common Objects in Context. Cocodataset.org. <http://cocodataset.org/#home\\> (2018)
22. Girshick, R., Donahue, J., Darrell, T., Malik, J.: Region-based convolutional networks for accurate object detection and segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **38**(1), 142–158 (2016)
23. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9 (2015)
24. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., Berg, A.C.: SSD: single shot multi-box detector. In: European Conference on Computer Vision, pp. 21–37. Springer, Cham (2016)
25. Teichman, A., Thrun, S.: Tracking-based semi-supervised learning. In: Robotics: Science and Systems (2011)
26. https://medium.com/@jonathan_hui/object-detection-speed-and-accuracy-comparison-faster-r-cnn-r-fcn-ssd-and-yolo-5425656ae359

Optimization of a Two-Cylinder Crankshaft by Computer-Aided Engineering



Basavaraj Talikoti, S. N. Kurbet and V. V. Kuppast

Abstract The most important aspect of an Internal Combustion (IC) engine is the crankshaft. A rotary motion is obtained from the crankshafts. Throughout its life, crankshaft is subjected to different levels of stress as it deals with the combustion of gases and load from the other parts like pistons and connecting rod; connected to. The pistons produce a reciprocating motion and are connected to crankshaft by the connecting rod. In order to maintain proper NVH levels and less air pollution by the combustion of gases, the primary requirement is to have a crankshaft with proper structure. If the crankshaft is unbalanced or has a degraded structure, due to prolonged use, then it leads to vibration and problems with the other parts that are connected to the crankshaft and which constitute the engine. Hence, it is a necessity to monitor the crankshaft for proper structure by optimizing it for stress at different conditions, so that the crankshaft has the best possible structure for a particular engine or vehicle where it is used.

Keywords Crankshaft · ANSYS · Stress · Computer-aided engineering (CAE)

1 Introduction

Crankshaft being an important facet of the automobile industry needs to be constantly upgraded to meet the demands of this progressive era. It is that component of the internal combustion engine which produces rotational motion. It basically converts the reciprocating motion of the pistons to rotational motion. There is a tremendous amount of stress acting on the crankshaft during its operation due to the various

B. Talikoti (✉)

Mechanical Engineering Department, Dr. K. M. Vasudevan Pillai College of Engineering Media Studies and Research, New Panvel, Maharashtra, India
e-mail: btalikoti21@gmail.com; blalilolt21@qrail.Cora

S. N. Kurbet · V. V. Kuppast

Mechanical Engineering Department, Basveshwar College of Engineering, Bagalkot, Karnataka, India

components attached to it [1–3]. The components attached to the crankshaft are the connecting rods, the pistons, flywheel, crankshaft balance masses (if required), and even the oil in the oil passages put some weight on the crankshaft. It has to endure the load from these parts connected to it and produce a smooth motion. With all these conditions, the crankshaft should be less noisy and should not produce much vibrations. Every kind of object; whether it be a construction, bridge, or even a vehicle at some point of time experiences random time variant vibrations or shocks which can be harmful. For example, when a vehicle passes an unnoticed pothole or bump on the road, it goes through random shocks or vibrations which can be harmful to it. In this way, a crankshaft also vibrates after long use, due to bending or an unbalance in its structure. Sometimes, the crankshaft can undergo extremely harmful torsional vibrations or a combination of different types of vibrations, which can twist and break the crankshaft. Thus, it is necessary to optimize the structure of the crankshaft in such a way that it will undergo minimum amount of stress in the most harsh conditions and produce none or less amount of vibrations and noise [4–6].

2 Optimization Technique by Analysis in ANSYS

To develop an optimized structure of crankshaft that will produce minimum stress in its structure in any operating conditions. As the production of crankshafts is in large scale and requires frequent improvements according to the needs of different applications or vehicles where it is used; the optimization process should be simple, quick, and accurate. Hence, the structure of the crankshaft is optimized virtually by using ANSYS software package. A three-dimensional structure of the crankshaft is designed and imported in ANSYS, after which the structure is simulated under different amounts of pressure that it can undergo in actual vehicle operation. Then, this structure is analyzed and the structure is modified to produce an optimized structure.

2.1 Importing the Geometry

The 3D model of the crankshaft is imported in ANSYS for analysis. After this, all the structural details of the crankshaft should be indicated in the model, to obtain accurate results. This is a very important step, as it where the analysis environment is programmed in such a way that the crankshaft behaves the way it will when it actually operates in the IC engine. Therefore, the geometry in ANSYS should be supported with proper settings like the fixed and flexible supports, the points in the structure where the crankshaft experiences load when it actually operates or where stress can result in deteriorating the crankshaft structure (Figs. 1 and 2).

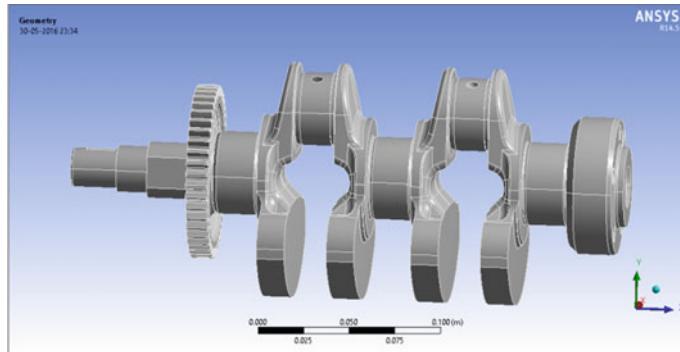


Fig. 1 Crankshaft geometry

Object Name	Fixed Support	Fixed Support 2	Fixed Support 3	Fixed Support 4	Force
State	Fully Defined				
Scope					
Scoping Method	Geometry Selection				
Geometry	146 Faces	88 Faces	93 Faces	463 Faces	6 Faces
Type	Definition				Force
Suppressed	Fixed Support				No
Define By					Vector
Magnitude					50678 N (ramped)
Direction					Defined

Fig. 2 Structural analysis settings in ANSYS workbench

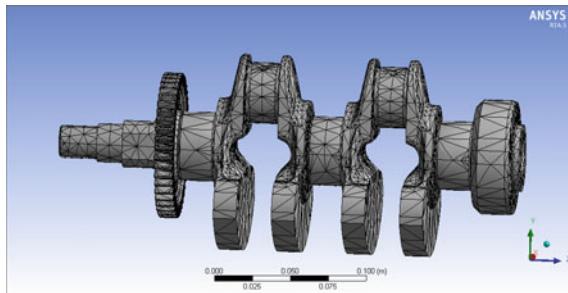
After giving the proper analysis conditions, meshing is performed. This is based on Finite Element Analysis (FEA). The entire structure is divided into a mesh. The structure is now composed of a finite number of elements and nodes. The fundamental of FEA is to breakdown a structure into a finite number of elements and analyze every element independently, here every mesh is analyzed separately. This particular crankshaft structure after meshing splits into 28874 elements and 50281 nodes (Fig. 3).

When performing meshing, care should be taken so that proper shape of the mesh element is chosen like triangular, tetrahedral, etc. The choice of shapes influences the accuracy of meshing. It should be seen that every part of the structure is meshed, and thus every element of the structure is analyzed (Fig. 4).

Object Name	Part 1		
State	Meshed		
Graphics Properties			
Visible	Yes	Suppressed	No
Transparency	1	Stiffness Behavior	Flexible
Bounding Box			
Length X	0.1074 m	Coordinate System	Default Coordinate System
Length Y	0.129 m	Reference Temperature	By Environment
Length Z	0.3061 m	Material	
Properties			
Volume	8.5564e-004 m ³	Assignment	Structural Steel
Mass	6.7167 kg	Nonlinear Effects	Yes
Centroid X	-2.8092e-005 m	Thermal Strain Effects	Yes
Centroid Y	-6.6935e-003 m	Density	7850 kg m ⁻³
Centroid Z	-0.12128 m	Coefficient of Thermal Expansion	1.2e-005 C ⁻¹
Moment of Inertia I _{p1}	3.963e-002 kg·m ²	Specific Heat	434 J kg ⁻¹ C ⁻¹
Moment of Inertia I _{p2}	3.6807e-002 kg·m ²	Thermal Conductivity	60.5 W m ⁻¹ C ⁻¹
Moment of Inertia I _{p3}	8.1579e-003 kg·m ²	Resistivity	1.7e-007 ohm m
Statistics			
Nodes	50281	Compressive Yield Strength Pa	2.5e+008
Elements	28874	Tensile Yield Strength Pa	2.5e+008
		Young's Modulus Pa	2.e+011
		Bulk Modulus Pa	1.6667e+011
		Shear Modulus Pa	7.6923e+010

Fig. 3 Specifications of the meshed structure defined and as seen in ANSYS workbench

Fig. 4 Meshed structure of the crankshaft



2.2 Static Analysis

After the boundary conditions are applied and meshing is done, the structure is evaluated for static loading conditions. This analysis provides the values for equivalent stress, total deformation, and total deformation. Equivalent stress which can also be termed as the von Mises stress here, is the value of stress which decides whether the structure will fail or it will be strong enough to sustain the load and operate under any condition that it can encounter while functioning. If the von Mises stress is less than the Yield stress of the material, then the structure will breakdown. The harmful frequencies of vibration which lead to other vibrations and can break the crankshaft can also be found out by static analysis (Figs. 5 and 6).

The results obtained in static structural analysis indicate that maximum equivalent stress is seen at the point in the structure where the crankshaft is connected to the

Fig. 5 The parts of the structure which are act as fixed supports or flexible supports are indicated in the boundary conditions

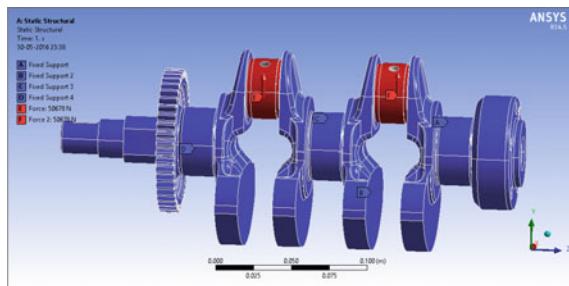


Fig. 6 Areas of maximum stress (equivalent stress result)

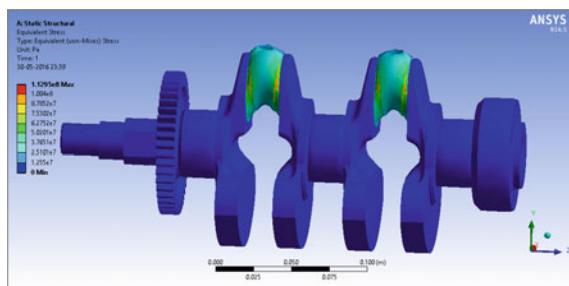
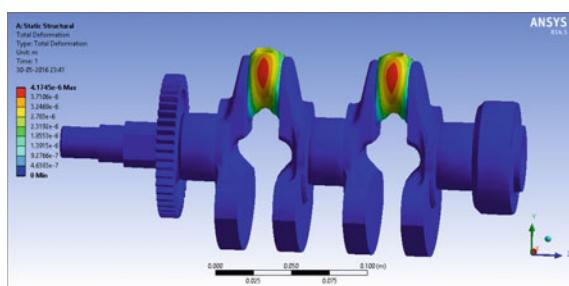


Fig. 7 Areas of maximum deformation (total deformation)



connecting rods which connect the pistons to the crankshaft. As maximum weight of the components connected to the crankshaft is given by the pistons and the connecting rods. Hence, the location of the structure at which maximum stress ($1.1295e + 008$ Pa) is observed is justified (Fig. 7).

Also, maximum amount of deformation is seen at the center of the crankpin. Thus, by static structural analysis; the areas of maximum stress and deformation can be found out.

2.3 Optimization

As it is seen in the analysis that some amount of stress is present in the crankshaft. Optimization is performed by removing and adding some weight to the crankshaft structure, so that it can withstand the pressure and the amount of stress observed is reduced.

3 Results

Figure 8 shows the optimization of crankshaft, horizontal axis represents the weights, the positive side is the addition of weight and the negative side is the removal of weight, the vertical axis indicates the stress level in all the trials. It is seen that the left side of the graph shows the removal of weight and the right side shows addition of weight, the vertical axis shows the stress observed in the crankshaft for addition and removal of different amount of weights. It is seen that in both the cases, the stress is above 0 Pa, which indicates either the removal or the addition of weights to the crankshaft is further increasing stress. Hence, it can be said that the crankshaft is optimized.

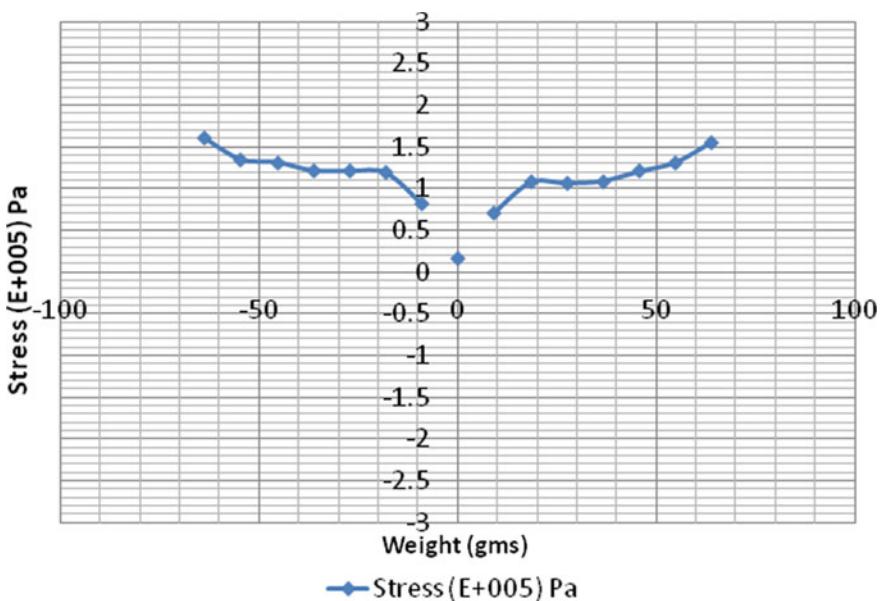


Fig. 8 Optimization of standard Tata Nano crankshaft

4 Conclusion

Thus, a technique for the optimization of the crankshaft structure has been developed by using Computer-Aided Engineering, which consumes less amount of time as compared to the regular techniques and also gives accurate results.

References

1. Mourelatos, Z.P.: A Crankshaft System Model for Structural Dynamic Analysis of Internal Combustion Engines. Vehicle Analysis and Dynamics Lab, Elsevier Science Ltd. (2001)
2. Crankshaft from an inline four cylinder engine with pistons, connecting rods and flywheel, Illustration copyright Eaglemoss publications/Car Care Magazine
3. Tata Motors Limited, Mahindra and Mahindra Ltd
4. Shah, P.D., Bhabhor, K.K.: Parametric optimization of four cylinder engine crankshafts. *Int. J. Eng. Sci. Inven.* **3**, 38–43 (2014)
5. Thriveni, K., Jaya Chandraiah, B.: Modeling and analysis of the crankshaft using ANSYS software. *Int. J. Comput. Eng. Res.* **3**, 84–88 (2013)
6. Rao, S.S.: Mechanical Vibrations, 5th edn. Pearson Education, Inc., publishing as Prentice Hall (2004) (Copyright 2011)
7. Homik, W.: Diagnostics, maintenance and regeneration of torsional vibration dampers for crankshafts of ship diesel engines. *Pol. Marit. Res.* **17**(1), 62–68 (2010)
8. Yingkui, G., Zhibo, Z.: Strength analysis of diesel engine crankshaft based on PRO/E and ANSYS. In: Third International Conference on Measuring Technology and Mechatronics Automation (2011)
9. Vijaykumar, K., Manoj, U., Girish, T., Rajesh, M., Girish, S.: Crankshaft design optimization to improve dynamic balancing and fatigue strength. *Int. J. Automot. Eng. Soc. Automot. Eng. Jpn.* 59–66 (2015)
10. Talikoti, B., Kurbet, S.N., Kuppast, V.V., Yadwad, A.M.: Modal analysis of a 2-cylinder crankshaft using ANSYS. *Int. J. Eng. Res. Appl.* (2015)
11. Talikoti, B.S., Kurbet, S.N., Kuppast, V.V.: Static structural analysis of 2-cylinder crankshaft using ANSYS. *Int. J. Technol. Res. Eng.* **3** (2015)
12. Talikoti, B.S., Kurbet, S..N., Kuppast, V.V.: A review on vibration analysis of crankshaft of internal combustion engine. *Int. Res. J. Eng. Technol. (IRJET)* **02** (2015). (e-ISSN: 2395 -0056)
13. Talikoti, B.S., Kurbet, S.N., Kuppast, V.V.: Harmonic analysis of a two cylinder crankshaft using ANSYS. In: ICICT. IEEE (2016)

Reducing Ping-Pong Effect in Heterogeneous Wireless Networks Using Machine Learning



Pragati Kene and S. L. Haridas

Abstract Ping-pong effect is inherent in handoff ready networks. Due to this effect, there are unnecessary handoffs in a wireless network, which reduces the energy efficiency and the quality of service (QoS) offered by the networks to the node under study. Ping-pong effect is also responsible for high delays in networks, due to the fact that the nodes and networks are constantly performing handoff operations, and thus the normal communication operations get delayed. This paper throws light on a novel cloud-oriented machine learning based approach which optimizes the number of handoffs in the network, and thus improves the overall QoS delivered by the network to the node. Our results imply that the ping-pong effect reduces by more than 20% when compared to standard handoff algorithms.

Keywords Ping-pong · Handoff · Heterogeneous networks · Delay · QoS

1 Introduction

Handoff in wireless networks basically means to shift nodes from the current low QoS network, into a better QoS network. The QoS of the network can be determined by the bandwidth provided by the network, the signal quality of the network, the cost per bit incurred while using the network, the data rate supported and many other network and node-based features. Based on these parameters, the nodes form a shifting criteria, which decides whether to shift the node to the new network or to keep itself in the current network. The shifting criteria is also called as handoff criteria, and is the single most important feature while developing an algorithm for handoff.

P. Kene (✉)

Department of Electronics Engineering, G.H. Raisoni College of Engineering, Nagpur, India
e-mail: pragati.kene@gmail.com

S. L. Haridas

Department of Electronics & Telecommunication Engineering, G.H. Raisoni
College of Engineering, Nagpur, India

A handoff criteria is weak, if it keeps on switching nodes from network to network very frequently, or, if it doesn't shift the node to required network even though the current network QoS has degraded to a very low level. Weak handoff criterions produces an effect in the network, which is called as the ping-pong effect. Due to this effect, the node shifts from network to network very rapidly, and usually doesn't stay in one network for an optimal period. Due to this effect, the following issues are faced by the node and the network,

- Reduced node lifetime

Due to increase in number of handoffs, the node is constantly shifting its data between different networks. Due to this fact, there is a lot of information transfer between the controlling network infrastructure and the node itself. This causes unwanted transmissions from the node to the network, and thereby degrades the energy of the node, thus affecting the node lifetime. Due to reduction in node lifetime, there is a straight effect on the course of life of the network itself.

- Increased node to node communication delay

As the node is not consistently maintaining its registration in one network, therefore there will be a continuous change in the QoS supplied by the network to the node, this changing QoS reduces the speed at which data transfers are being done in the network, which increases the jitter of the network. Usually an increase in jitter, also increases the average node to node communication delay in the network.

- Reduction in network serving speed

Due to ping-pong effect, the network is being continuously pinged by the nodes, thus a notable amount of network processing time is being utilized for performing handoffs. Due to this, the network gets lesser count of clock cycles for performing actual communications, thus there is a notable reduction in the network serving speed, which inherently reduces the QoS of the network.

- Reduction in communication throughput

Due to reduction in network serving speed, there is an increase in communication delay within the network. This increase in communication delay, causes the network's throughput to reduce. Reduction in throughput again affects the bandwidth provided by the network, and thereby reduces the QoS of the network.

Due to these issues, the network and the node recursively reduce the performance of each other, and thereby the overall network quality is reduced. Thus, the ping-pong effect is a very dangerous condition which must be carefully sorted out by proper planning of the handoff criterion. Here, we proposed a machine learning oriented algorithm, which uses real-time training and testing data for the purpose to evaluate the best handoff criterion for the network. This criterion depends upon a vector of parameters, which are carefully selected so that the ping-pong effect is reduced, thereby optimizing the number of handoffs in the network. Result evaluation of our proposed algorithm, and comparison with the standard handoff criterion methods indicate that our proposed algorithm has lower number of handoffs, while maintaining a high QoS and thereby a reduced ping-pong effect.

2 Literature Review

Numerous calculations have been studied to decrease pointless handoffs in systems. Like, in [1] the analysts suggest that Ultra-thick arrangement of little or tiny cells (SC) can be anticipated in 5G, organize under the inclusion region of the Macrocell (MC). A versatile client (UE) ought to have the capacity to find nearby SCs to play out the handover (HO). This procedure should be possible by incessant neighbor cell checking. Nonetheless, broad filtering for each SC in a thick organization situation is an asset squandering methodology, which results in power dispersal of the UE battery and furthermore brings down the throughput gain. This likewise implies that high number of SCs would be accessible for the UE to carry HO to. Henceforth, the likelihood of pointless HO will increment and thus debase the UE's nature of administration (QoS). In this work, the specialists plan to limit pointless HOs in two level heterogeneous system with thick organization of SCs. Output illustrates that the proposed calculation beat the ordinary HO strategy with decreased superfluous HOs and expanded throughput for the system especially for medium to rapid UEs bringing about great UE QoS. While in [2], the specialists propose that ideal settings for Handover parameters (Hysteresis and Time to Trigger) rely upon client speeds in the system. The self-organization networks (SON) standard characterizes the mobility robustness optimization (MRO), and utilizes it in case for the independent strategies for arranging the parameters in coinciding to the portability of design. Best in class MRO arrangements have depended on master learning, control-based calculations to look through the parameter space; yet it is clumsy to configure rules for all conceivable portability designs in any system. In this work, the specialists present a Q-learning MRO arrangement, QMRO, which takes in the required parameter esteems, proper for particular speed conditions in the individual cells. They think about QMRO against the best static reference arrangement (Ref) that is acquired by sweeping the parameter space. Their outcomes demonstrate that QMRO can learn parameter settings that accomplish comparable execution to Ref in a practical system condition where clients have powerful movement models. In this way machine learning can be utilized for enhancement of handoff process.

In [3], the scientists have taken a shot at the cutting edge remote systems which give heterogeneous remote access. This heterogeneous system gives consistent availability and in every case best-associated administrations. In their work, organize parameters are utilized to channel the competitor arrange set. Rules-based handoff calculation is contrasted and proposed calculation and the reproduction results imply that the proposed calculation's execution is upgraded by decreasing superfluous handoffs.

In [4], the analysts recommend that handover is one of the key activities in the portability administration of long haul advancement (LTE)-based frameworks. Hard handover chosen by handover edge and time to trigger (TTT) has been received in 3G Partnership Project (3GPP) LTE with the motivation behind decreasing the unpredictability of system engineering. Different handover calculations, in any case, have been focused for 3GPP LTE to boost the framework goodput and limit packet

communication delay. In [4], another handover approach upgrading the current handover plans is proposed. It is significantly dependent on the two ideas of handover administration: lethargic handover for abstaining from ping-pong impact and early handover for dealing with continuous administrations. Lethargic handover is bolstered by forbidding handover before the TTT window lapses, while early handover is upheld even before the window terminates if the rate of change in flag control is huge. The execution of the proposed plan is assessed and contrasted the handover calculations dependent on goodput per cell, normal parcel delay, number of handovers every second, and flag to obstruction in addition to commotion proportion. Work in [5] utilizes delicate figuring procedures like ant colony optimization (ACO) to decrease ping-pong impact. The work in [5], proposes another methodology in which mobile users (MTs) ceaselessly dissect the system and keep a database of the best accessible systems. The idea depends on QoS mindful insect settlement based on the vertical handoff component that uses a refreshed approach of ants' province streamlining decision calculation (ACOR), using the dynamic and static factors, for example, RSS, the expense of administration, data transfer capacity, the speed of MT, the power utilization and security, and the module for foreseeing the voyaging separation inside an IEEE 802.11 WLAN cell. Recreation results in [5], demonstrate that we can not just meet the individual needs of clients as far as QoS is considered the entire framework execution by diminishing the quantity of handover disappointments and superfluous handover examples.

Work in [6], utilizes broad geometric and probabilistic procedures to build up a practical and novel model for the inclusion territory of a remote neighborhood (WLAN) cell with a plan to limit pointless handover and handover disappointment of a nomadic hub (MN) navigating the WLAN cell. The abide time is assessed alongside the edge esteems to guarantee an ideal handover decision by the MN, while the likelihood of superfluous handover and likelihood of handover disappointment are kept inside middle of the road limits. Monte Carlo reenactments were done to evaluate the conduct of the proposed and existing models. Reproduction results in [6], demonstrated that the proposed model is more strong and fit for keeping superfluous handover likelihood and handover disappointment likelihood nearest to a predefined likelihood benchmark.

The plan proposed in [7], is a streamlined handover decision system, “Voyaging Time Prediction Based on the Consecutive RSS Measurements” in an IEEE 802.11 WLAN cell is recommended. The strategy utilizes a period edge which is figured by a Mobile Terminal (MT) when it infiltrates into a WLAN limit. The assessed voyaging time is then contrasted and the time limit, in order to settle on handover decisions for lessening the likelihood of handover disappointments. The execution examination on [7], uncovers that the recommended system successfully limits the quantity of handover disappointments by 60% when contrasted with the as of now proposed plans.

3 Proposed Handoff Technique

Optimization of network QoS by reducing the number of handoffs is the main intention of this research. This problem is complicated enough, and thus needs a machine learning oriented optimization algorithm to perform the task. In this section we describe our machine learning based algorithm which uses QoS maximization along with number of handoff minimization using a series of iterative steps. Majorly our algorithm has to take a decision on the number of parameters to be used for handoff, and then use a score based threshold approach to perform handoff decision. The algorithm performs the handoffs using the following steps:

- Initialize the algorithm parameters,

N_{max} = Max no of parameters

NR = No of Rounds

N_s = No of Solutions

$LTHSF$ = Learning Threshold Speed factor range is 0–1

- Step1: Select “K” no of random parameter
- Step 2: Perform Network handoff and obtain
 - (a) No of Handoff = NH
 - (b) Handoff Delay = DH
 - (c) Throughput = TH
- Step 3: Evaluate the learning factor by using

$$L.F = \frac{TH}{(NH + DH)} \quad (1)$$

- Step 4: For N_s solution, we get N_s values of $L.F$ from $L.F_1$ to $L.F_{N_s}$
- Step 5: Evaluate Mean of $L.F$

$$L.FMean = \left(\sum_{i=1}^{N_s} (L.F_i) \right) / N_s$$

- Step 6: Evaluate the Learning Threshold

$$LTH = L.FMean * LTHSF$$

- Step 7: For all the solutions where $L.F > LTH$ store the details of that solution such as number of parameters, NH , DH , $L.F$ else discard that solution.
 - Step 8: Repeat the process for all the NR rounds to get the trained Handoff system
 - Step 9: At the last NR round, select the solution with maximum value of LF for Executing Handoff in the network
- Algorithm to further reduce Number of Handoff

If connection time increases, Number of Handoff decreases, Handoff failure decreases, also connection breakdown decreases.

1. If Handoff score > Threshold, then no need to handoff so, connection time $T_c = T_c + 1$;
2. Else if Handoff score < Threshold, then first calculate mean T_c if Connection time for particular node $T_{ci} >$ mean T_c and Handoff Score < Threshold/2 then Handoff is done Time for which node was in that network $T_{ni} = T_{ci}$;

$$T_{ci} = 0;$$

3. Else No handoff is required so $T_{ci} = T_{ci} + 1$;
4. If $T_{ci} > T_{ni}$ then $T_{ni} = T_{ci}$;

Our algorithm is flexible with respect to the number of parameters which can be used for handoff, and the underlying handoff decision algorithm, but for this text we have used Received signal strength indicator, Signal to noise ratio, Network bandwidth, Data rate, Network coverage, Node's interest in the given network, Link quality, Cost per byte for the network, Urgency of data transfer by the node, Max data size permissible by the network, and Coverage to speed ratio.

We used a score based handoff algorithm, which compares the value of the given parameter against a range as specified by the network, scores the network based on this comparison, and then selects the network which gets the highest score. This is standard game theory technique, and is used by various researchers for low complexity handoffs.

Due to the inclusion of the parameter N_h (number of handoffs) in Eq. 1, along with TH and DH , the algorithm selects a solution which has the smallest value of N_h and DH in the network, which can be obtained by keeping a sufficiently high value of TH . This is possible due to the fact that the learning factor (L_f) has N_h and DH as denominator terms, while TH is a numerator term, thus, the machine learning optimization maximizes the throughput, and reduces the delay and amount of handoffs for any selected solution to be used. This property of the aimed machine learning algorithm makes it very powerful, and helps in reducing the ping-pong effect from the network. The detailed analysis of these parameters is given in the upcoming section.

4 Performance Analysis

We are using network simulator version 2.29 for performance analysis of the developed handoff system. The multi-handoff criteria (MIH) patch has been used while developing the system, and has been integrated with the ns2.29 installation. Once the system parameters are setup, at that point we tried the framework dependent

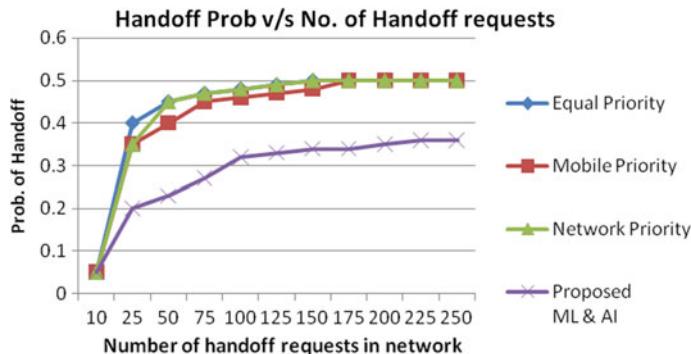


Fig. 1 Probability of handoff versus handoff requests

on particular scopes of RSSI, SNR, data rate and transfer speed for the WiFi and WiMAX systems. We at that point changed the quantity of nodes, did a few varieties in the count of movements and communications, and then analyzed the effect of our algorithm on probability of occurrence of handoff.

We observed that the proposed protocol reduces the probability of handoff to other networks when compared to standard protocols. This evaluation suggests that the proposed protocol reduces the ping-pong effect by 32% on an average, and thereby improves the overall efficiency of the network. The graphical image can be visualized from Fig. 1. The analysis shows that the existing algorithms saturate at about 100 handoff requests, while our algorithm tries to resist the change in count of handoff requests adaptively and does not saturate even for higher number of handoff requests. This ensures that the accomplishment of the algorithm remains optimum even under increasing number of nodes, which directly relates to number of handoff requests.

It can be said that as the node velocity increases, the number of handoffs over a fixed period of simulation time reduces as the node movement does not guarantee its network location at high speeds. From the Fig. 2, we can note that our algorithm lessens the number of handoffs by more than 10%, and therefore reduces the ping-pong effect when compared to other standard algorithms.

These results focus the effectiveness of our technique when compared to standard state of the affairs algorithms. We also checked the effect of reduction in count of handoffs and reduction of probability of handoffs on other node QoS parameters like end-to-end delay, network throughput, energy consumption, and packet delivery ratio, and observed that these parameters do not vary with reduction in number of handoffs as compared to other algorithms, thus indicating that the other QoS parameters are not affected due to the machine learning optimization.

The values of those parameters vary with network and node configurations, and thus is not mentioned in this text. Due to these advantages, this algorithm has added value that helps it to be used in real time for widely diversified networks, and varying network conditions.

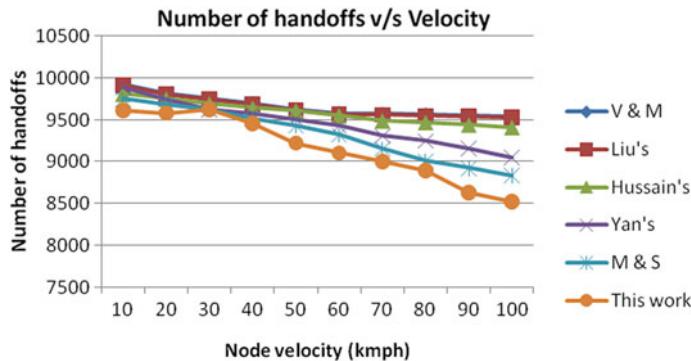


Fig. 2 Node velocity versus number of handoffs

5 Conclusion

The effect of machine learning-based optimization of the system is quite inherent. But the results show a major improvement in ping-pong effect removal from the network, and thereby making the proposed algorithm as a choice for widely diversified network configurations. Results show a reduction in handoff probability of 32% and a reduction in number of handoffs by more than 10% when compared to standard algorithms, which can be used as a base point for other algorithms in future comparisons.

Furthermore, researchers can work on security of the proposed machine learning-based handoff system, and try to integrate block chains and other state of the affairs encryption techniques in view of make the system efficient enough in terms of security of the network.

References

1. Alhabo, M., Zhang, L.: Unnecessary handover minimization in two-tier heterogeneous networks. In: 13th Annual Conference on Wireless On-demand Network Systems and Services (WONS), Wyoming, USA, pp. 160–164. International Federation for Information Processing (2017)
2. Mwanje, S.S., Mitschele-Thiel, A.: Distributed cooperative Q-learning for mobility sensitive handover optimization in LTE SON. In: IEEE Symposium on Computer and Communications (ISCC) (2014)
3. Kulkarni, M.P., Bale, V.S.: Vertical handoff decision algorithm to reduce unnecessary handoff for heterogeneous wireless networks. (*IJCSIT*) Int. J. Comput. Sci. Inf. Technol. **6**(6), 5033–5036 (2015)
4. Chen, X., Tae Kim, K., Lee, B., Youn, H.Y.: DIHAT: differential integrator handover algorithm with TTT window for LTE-based systems. *EURASIP J. Wirel. Commun. Netw.* (Springer) (2014)

5. El Fachtali, I., Saadane, R., ElKoutbi, M.: Vertical handover decision algorithm using ants' colonies for 4G heterogeneous wireless networks. *J. Comput. Netw. Commun.* (Hindawi Publishing Corporation) **2016**, Article ID 6259802, 15 (2016)
6. Omoniwal, B., Hussain, R., Ahmed, J., Iqbal, A., Murkaz, A., Ul-Hasan, Q., Malik, S.A.: A novel model for minimizing unnecessary handover in heterogeneous networks. *Turk. J. Electr. Eng. Comput. Sci.* **26**, 1771–1782 (2018)
7. Mahmood, A., Zen, H., Othman, A.K., Siddiqui, S.A.: An optimized travelling time estimation mechanism for minimizing handover failures from cellular networks to WLANs. In: 2015 International Conference on Estimation, Detection and Information Fusion (ICEDIF 2015), pp. 28–33 (2015)

Design and Analysis of Elliptic Curve Cryptography-Based Multi-round Authentication Protocols for Resource-Constrained Devices



Adarsh Kumar, Alok Aggarwal, Neelu Jyoti Ahuja and Ravi Singhal

Abstract Authentication is among the initial steps of network formulation. Lightweight authentication processes are preferred for resource-constrained devices. These processes are efficient in terms of computational and communicational costs but inappropriate in providing required security rating. This work explores the feasibility of elliptic curve-based cryptography (ECC) for authentication processes. Three lightweight authentication protocols, which use ECC are identified and analysed for hierarchical network. These protocols are simulated and validated using lightweight cryptography primitives and protocols. In results, it is observed that out of three protocols protocol 1 performs better for 75 and 1000 nodes networks. In comparative analysis, a minimum of 14.1% and maximum of 24.7% throughput improvement is observed. Further, delay is reduced by 15.6% (minimum) and 31.3% (maximum) for protocol 1 as compared to protocol 2 and protocol 3.

Keywords ECC · Authentication · RFID · Reader · Tag · Simulation · QoS

1 Introduction

Internet of Things (IoT), consisting of RFID and sensor devices, is expanded in various application domains like [1] health care, supply chain management, traffic control systems, waste management systems, security primitives and protocols, etc. Mobile ad hoc Networks (MANETs) added self-configuration, autonomous identifi-

A. Kumar (✉) · A. Aggarwal · N. J. Ahuja · R. Singhal
University of Petroleum and Energy Studies, Dehradun, India
e-mail: adarsh.kumar@ddn.upes.ac.in

A. Aggarwal
e-mail: alok.aggarwal@ddn.upes.ac.in

N. J. Ahuja
e-mail: neelu@ddn.upes.ac.in

R. Singhal
e-mail: r.singhal@ddn.upes.ac.in

cation, authentication and routing and decentralized network to IoT networks. Using RFID in IoT, objects are uniquely identified initially for subsequent processing [2]. After identification, MANET node features help in information exchange and data transmission over internet. Thus, MANET, RFID and Sensor nodes integrated network constitutes IoT network. This way, a network of any object can be constituted for information gathering and gain the support system for controlling the processes [3]. Further, a network may consists of resourceful or resource-constrained devices. Resourceful devices are capable enough for handling any computational and communicational complexities for providing maximum security. Resource-constrained devices requires efficient mechanisms for providing higher security. In order to integrate cryptography primitives like confidentiality, integrity, authentication, availability and non-repudiation, more hardware resources or hardware efficient mechanisms are required.

In authentication protocols are classified based on resourceful and resource-constrained devices. Resourceful devices use cryptographic primitives having higher computational and communication costs. Whereas, resource-constrained devices prefer to use lightweight cryptographic primitives [4], or least computation and communication cost processes for authentication. In resource-constrained devices, hardware cost of cryptography primitives and protocols is minimized by measuring the gate equivalents (GEs) [5–7]. Apart from lightweight and regular authentication protocols, ultra-lightweight authentication process uses logical operations at each end for challenge verification and authentication. These processes consume least computations in identification, authentication and record management. Further, these processes are classified either single or multi-round authentication processes. Challenging side sends a challenge to other side and receiver side has to verify the challenge for authentication. If verification side's answer is verified, then object is authenticated.

In RFID systems, three parties help in authentication process. These parties are reader, tag and back-end data centre. Further, it is analysed that there are following possible scenarios of lightweight authentication in resource-constrained devices. Figure 1 shows authentication mechanisms with single round without any functional or polynomial computations (scenario 1). Figure 2 shows single-round authentication process with functional or polynomial computations (scenario 2). Figure 3 shows multi-round authentication process with functional or polynomial computations.

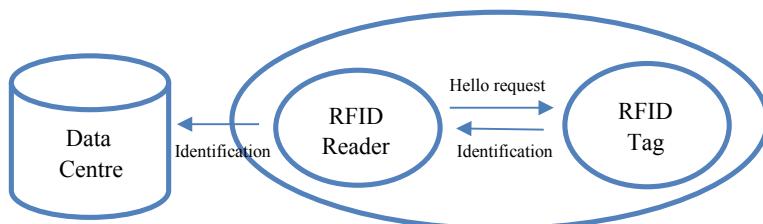


Fig. 1 Single round without functional/polynomial lightweight authentication (Scenario-1)

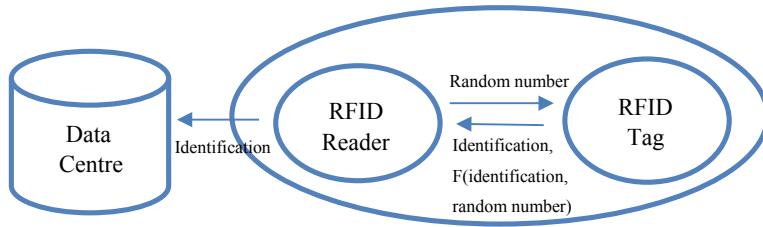


Fig. 2 Single round with functional/polynomial lightweight authentication (Scenario 2)

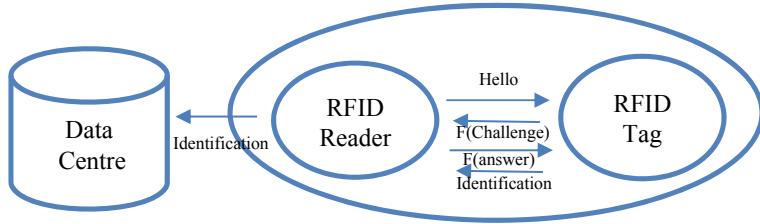


Fig. 3 Multi-round with functional/polynomial lightweight authentication (Scenario 3)

This work analyses three asymmetric key cryptosystem-based authentication protocols for network consisting of resource-constrained devices. These devices include RFID reader, RFID tag and sensor nodes. A network consisting of these devices are attached to objects for their authentication and their functionality in the network. A centralized network of interconnection of clusters is formalized for analysis. Each object in this network is part of cluster and network, and acts as a smart object. A smart object is an object having autonomous identification, routing and data exchange feature over longer range. This work ensures authentication of smart network objects through elliptic curve-based cryptosystem. In performance analysis, a network of 75–1000 nodes is constructed for QoS measurements. A comparative analysis of three protocols is performed using throughput and jitter QoS parameters.

Organization of work is as follows: state-of-the art survey is performed in Sect. 2. This survey analyses the authentication protocols for resource-constrained devices in asymmetric key cryptosystem category. In Sect. 3, three asymmetric cryptography protocols are discussed with their mathematical computations in authentication process. Section 4 presents network construction process, integration of authentication protocol with hierarchical network, variation in number of nodes for performance analysis and QoS measurements. Section 5 concludes the work.

2 Literature Survey

A family of ultra-lightweight mutual authentication protocols is proposed in [9, 10] which were reported in [11] that these protocols are vulnerable to desynchronization attack and full-disclosure attack. For providing strong authentication and strong integrity another ultra-lightweight protocol called SASI is proposed in [12]. Vulnerabilities have always been a vital issue such as tag traceability, desynchronization and secret disclosure [13–15]. Work proposed in [12] suffered from many vulnerabilities like man-in-the-middle attack [13], tracking attacks [14], de-synchronization attack [15]. A new attack dedicated to ultra-lightweight protocols, Tango attack is proposed in [16]. A new ultra-lightweight RFID authentication protocol with permutation (RAPP) is proposed in [17] by introducing the permutation operation to break the orders of the bits. To resist de-synchronization attacks, both the old value and the new value of the shared keys are stored in the back-end database. By doing so the storage of the tag is also economized. It is shown that proposed RAPP achieves privacy and security and requires fewer resources for tags compared with other ultra-lightweight authentication protocols.

Based on chaotic maps in [18] an authenticated RFID security mechanism is proposed based on chaotic maps that guarantee a mutually authenticated process and uses a simple and robust chaotic cryptosystem to implement a practical RFID mechanism. Focus is on the security of the transmitted information and the privacy of personal information or data were. A detailed information is provided concerning authentication proof and security aspects of our new design. Authors in [19] have focused on some weaknesses in the SASI protocol and showed that the protocol presents vulnerabilities which can be easily used by an adversary who can interact with the Tag. Three attacks have been considered which built one on the top of the other; namely desynchronization attack, identity disclosure attack and full-disclosure attack. The design of a secure and privacy-preserving RFID authentication protocol requires an appropriate security and privacy model to enable a careful analysis of the protocol. Many models were proposed like in [20] a suitable model for a secure and privacy-preserving is proposed which was extended in [21]. In [22] a model based on UC framework is proposed. Model proposed in [23] focused on the design and the analysis of protocols based on symmetric-key encryption schemes, and on the trade-off security versus efficiency by adding completeness and soundness requirements. The authors in [24] focused on a model based on random oracles and indistinguishability.

3 Proposed ECC-Based Authentication Protocols

In this section, ECC-based lightweight authentication protocol's mathematics and functionalities are discussed in detail. Protocols and their functionalities are as follows:

Protocol 1: EC-RAC 2

Step 1: Tag ‘T’ generates a challenge with the help of random number ‘ e_1 ’ and base point on elliptic curve ‘E’. Tag sends this challenge to reader ‘R’.

$T \rightarrow R : e_1 P$

Step 2: Reader generates a random number ‘ r_1 ’ and sends it to tag.

$R \rightarrow T : r_1$

Step 3: Tag computes two challenges ‘Temp₁’ and ‘Temp₂’ with the help of random numbers (e_1 and r_1), password of tag stored at data centre (PASSWD^T), identification of tag (ID^T) and public key of reader (PU^R). Tag sends these challenges to reader.

$T : Temp_1 = (e_1 + r_1 ID^T) PU^R, Temp_2 = (e_1 ID^T + r_1 PASSWD^T). PU^R$

$T \rightarrow R : Temp_1, Temp_2$

Step 4: Reader verifies received challenges with the help of password verifier (PASSWD-VERIF^T), private key of reader (PR^R) and inverse operations. If challenge is verified then tag is considered to authentic else unauthentic.

$R : ((PR^R)^{-1} Temp_1 - e_1 P) r_1^{-1} = ID^T P$, Now find ID^T entry and extract PASSWD-VERIF^T.

: if $((PR^R)^{-1} Temp_2 - ID^T \cdot e_1 \cdot P) r_1^{-1}$ equals to PASSWD-VERIF^T then accept else reject.

Protocol 2: EC-RAC 3

Step 1: In the first step, tag ‘T’ computes two challenges for reader with random numbers e_1 and e_2 , and base point on elliptic curve ‘P’. Tag sends these challenges to reader.

$T \rightarrow R : e_1 P, e_2 P$

Step 2: Reader ‘R’ reply back with a random number to tag.

$R \rightarrow T : r_1$

Step 3: Tag generates two challenges with the help of tag selected random numbers (e_1 and e_2), reader selected random number (r_1), tag password stored in data centre (PASSWD^T), identification of tag (ID^T) and public key of reader (PU^R). Tag sends these challenges to reader.

$T : Temp_1 = (e_1 + r_1 ID^T) PU^R, Temp_2 = (e_2 ID^T + r_1 PASSWD^T). PU^R$

$T \rightarrow R : Temp_1, Temp_2$

Step 4: Reader verifies the received challenges with the help of its own private key (PR^R), random number (e_1 and r_1), password verifier (PASSWD-VERIF^T) and inverse operations. If challenges are verified then tag is considered to authentic else unauthentic.

R : $((PR^R)^{-1}Temp_1 - e_1 P) r_1^{-1} = ID^T P$, Now find ID^T entry and extract PASSWD-VERIF T .

: if $((PR^R)^{-1}Temp_2 - ID^T \cdot e_2 \cdot P) r_1^{-1}$ equals to PASSWD-VERIF T then accept else reject.

Protocol 3: ERAP (ECC-based RFID Authentication Protocol).

Step 1: Reader generates a random number challenge ‘ r_1 ’ and sends it to tag ‘T’.

R → T : r_1

Step 2: First, Tag computes a point ‘P’ using a new random number ‘ e_1 ’ and generator on elliptic curve ‘G’. Another coordinate (x_T , y_T) is calculated from (x_p , y_p) using ‘ e_1 ’, ‘ r_1 ’, private key of tag, i.e. $PR_T (=e_3)$ and inverse operations. Tag reply back with new coordinates (x_T , y_T) and new random number ‘ e_2 ’.

T : Compute $P = e_1 G = (x_p, y_p)$

: if $x_p \in F_n$ then $x_P^I \in [1, n - 1]$ else if $x_P^I \in F_{2^n}$ then $x_P^I = \sum_{i=0}^{n-1} 2^i x_p$

: Compute $x_T = x_P^I \bmod n$ and $y_T = e_1^{-1}(r_1 + PR_T \cdot x_T)$

: if x_T or y_T is zero then recalculate step 2.

T → R : (x_T , y_T) and e_2

Step 3: Initially, received coordinates are verified, i.e. whether x_T and $y_T \in [1, n - 1]$. If any of these coordinates is not verified then tag is considered to be unauthentic else authentic and continues. Reader uses $PU_T (=e_3 G)$, i.e. the public key of tag to generate a new challenge for tag (x_R , y_R). Reader sends this challenge to tag.

R : Compute $w = (y_T)^{-1} \bmod n$, $u_1 = r_1 w \bmod n$, $u_2 = x_T w \bmod n$ and $P' = u_1 G + u_2 \cdot PU_T$. If $P' = \infty$ then tag is considered to be unauthentic else continue.

: if $x_{P'} \in F_n$ then $x_{P'}^I \in [1, n - 1]$ else if $x_{P'}^I \in F_{2^n}$ then $x_{P'}^I = \sum_{i=0}^{n-1} 2^i x_{P'}$

: Now, if $x_T = x_{P'}^I \bmod n$ then tag is authentic else unauthentic.

: Authentic tag will compute $P'' = r_2 PU_T$ and if $x_{P''} \in F_n$ then $x_{P''}^I \in [1, n - 1]$ else if $x_{P''}^I \in F_{2^n}$ then $x_{P''}^I = \sum_{i=0}^{n-1} 2^i x_{P''}$, $x_R = x_{P''}^I \bmod n$, $y_R = r_2^{-1}(e_2 + PR_R x_R) \bmod n$, if x_R or y_R is zero then recomputed these variables by selecting another value of r_2 and computing P'' .

R → T : (x_R , y_R)

Step 4: Here, tag also verifies that whether x_R and $y_R \in [1, n - 1]$, if anyone is not verified then tag is considered to be unauthentic else authentic and continues. Tag computes a point P' using the seed used to randomly generate the elliptic curve, i.e. ‘S’. It checks the range of P' . If P' lies within acceptable range then tag is considered to be authentic else unauthentic.

T : Compute $w = (y_R)^{-1} \bmod n$, $u_1 = e_1 w \bmod n$, $u_2 = x_R w \bmod n$ and $P' = (u_1 + u_2 S) PR_T G$. if $P' = \infty$ then tag is considered to be unauthentic else continue.

: if $x_{P'} \in F_n$ then $x_{P'}^I \in [1, n - 1]$ else if $x_{P'}^I \in F_{2^n}$ then $x_{P'}^I = \sum_{i=0}^{n-1} 2^i x_{P'}$

: Now, if $x_R = x_{P'}^I \bmod n$ then tag is authentic else unauthentic.

Table 1 Simulation parameters

Parameters	Value
Channel type	Wireless channel
Radio propagation model	Two ray ground
Network interface	WirelessPhy
MAC type	802.11
Interface queue	Priority queue
Antenna	Omni antenna
Max packets in queue	100
Routing protocol	AODV
X dimension of the topography	800 m
Y dimension of the topography	800 m
Mobility model	Random waypoint mobility
Data rates	Three packets/s
Packet size	512 bits
Simulator	ns-3
Simulation time	1000 s

4 Results and Analysis

This section presents network setup, simulation parameters taken for analysis and description of results. Simulation analysis is performed over 50–1000 nodes network. Three network scenarios are taken for analysis. Networks of 50 and 150 nodes are considered in small-scale network category whereas a network of 1000 nodes is considered in large-scale network. Following is the description of simulation setup, results and its descriptions.

Simulation Parameters: Details of simulation parameters are shown in Table 1. Three authentication protocols are tested in a hierarchical network consisting of interconnection of clusters. These clusters are connected through their cluster heads in order to formalize a virtual hierarchical network. This virtual network consists of readers and tags. A reader is connected with cluster head to control nodes to which tags are attached for identification and verification.

Jitter: It is defined as the deviation in root mean square value of average delay. Minimum jitter value indicates the stability of network and increases in jitter values increases the chances of attack. In this work, a small-scale network of 75 and 150 nodes are considered for analysis with different packet delivery rates (0.1 pkt/s, 1 pkt/s and 5 pkts/s) and with different authentication protocols as shown in Figs. 4 and 5. Results show that protocol 3 is better for 75 nodes network and protocol 1 is better for 150 nodes network. Trends show that increase in number of nodes decreases jitter values because more nodes are available for packet transmission, which in turn reduces delay and its variations. Performance is further improved by increasing the packet delivery rate. A higher rate results in reducing delays and its

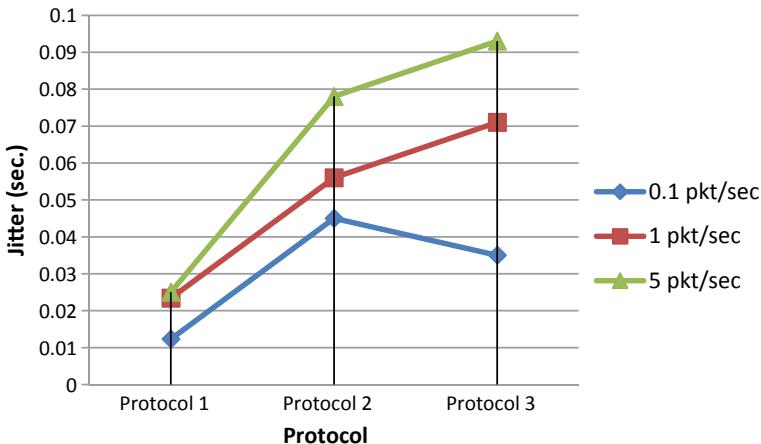


Fig. 4 Jitter analysis of three authentication protocols in a network of 75 nodes

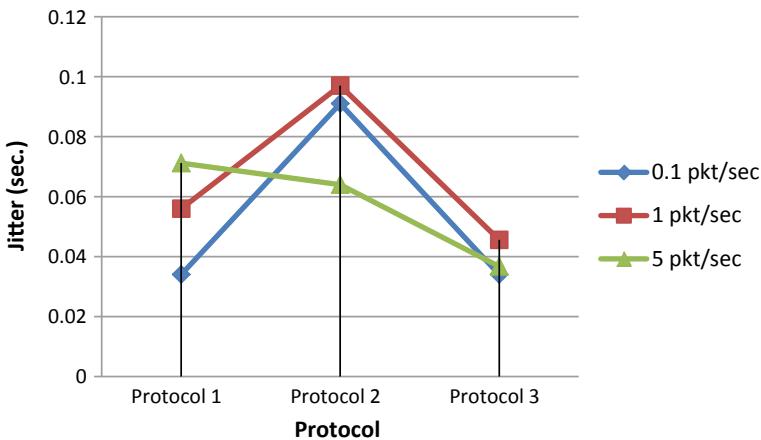


Fig. 5 Jitter analysis of three authentication protocols in a network of 150 nodes

deviations (jitter). Faster data rate leaves behind many empty routes for more data transmission which in-turn reduces the chances of attack by proper data transmission. Selection of routing protocol is also helpful in reducing jitter and chances of attacks. Routing protocol predicts the network packet drop rate and suggestion alternative routes immediately for reducing any types of delay.

Goodput: It is the ratio of number of packets delivered to its destination divided by the number of packets received for delivery. In goodput calculation, header information is not counted. Figures 6 and 7 show the comparative analysis of goodput variation over time for 75 and 150 nodes networks. It is observed that goodput and its variation is maximum for protocol 1 (for 75 nodes network) and for protocol 2 (for 150 nodes)

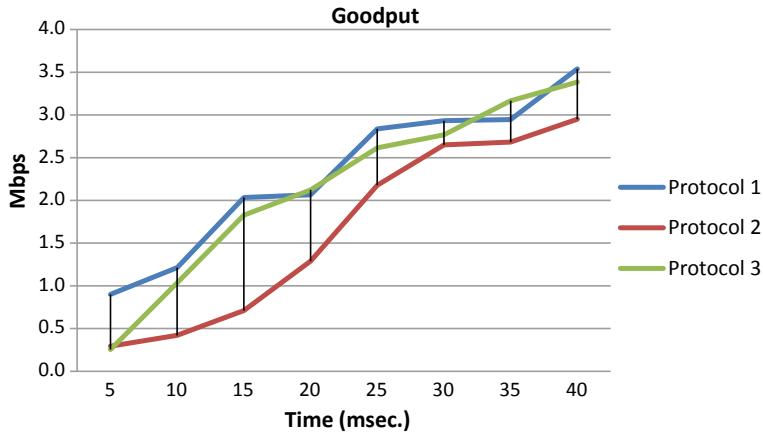


Fig. 6 Goodput analysis of three authentication protocols in a network of 75 nodes

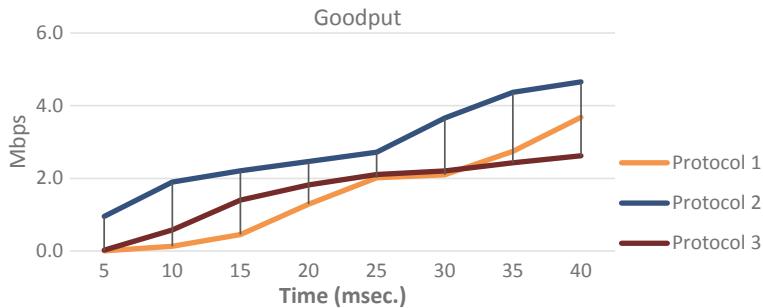


Fig. 7 Goodput analysis for three authentication protocols in a network of 150 nodes

network. Protocol 3's goodput is comparable to protocol 1 and protocol 2 but it is least as compared to protocol 1 and protocol 2. Goodput increases with increase in number of nodes (75–150 nodes) because more nodes are available for parallel route establishment. However, this increase is observed to be maximum for protocol 1.

In another analysis, performance is analysed for 1000 nodes network. Figure 8 shows the comparative analysis of three protocols for 1000 nodes network. In results, it is observed that protocol 3 is performing better as compared to other protocols. It is having minimum processing and jitter. This delay and jitter is minimum because of least computational and communicational costs. In other protocols, use of cryptography primitives and its scale is large, thus it is not performing good in processing the packet for transmission to next subsequent nodes. Although large number of nodes are available in 1000 nodes network for parallel route establishment but every node expects to establish a route for taking an advantage of being active node.

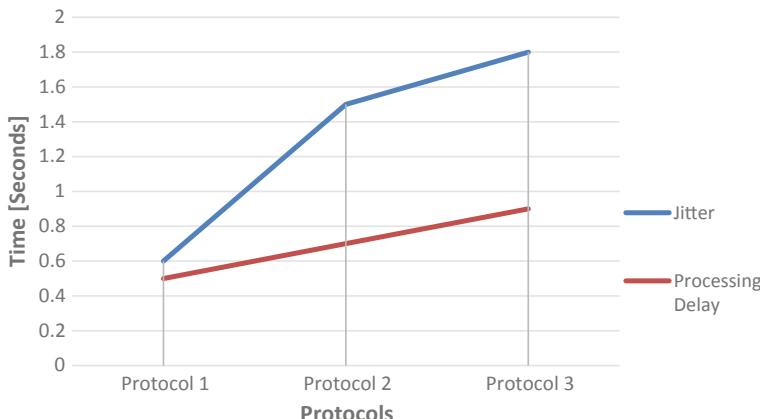


Fig. 8 Analysis of authentication protocols in a network of 1000 nodes

5 Conclusion

Authentication plays an important role in network construction and expansion. Protection through authentication increases with increase in complexity. However, lightweight authentication mechanisms are available for resource-constrained devices. This work is an initiative for analysing the asymmetric cryptography primitives for resource-constrained devices for achieving maximum security. In this work, three elliptic curve cryptosystem-based authentication protocols are analysed for resource-constrained devices in hierarchical MANET. A comparative analysis of QoS parameters is performed for 75–1000 nodes networks. Results shows that protocol 1 is performing better as compared to other protocols. It is giving higher throughput with minimum delay for 75 and 1000 nodes networks. A minimum of 14.1% and maximum of 24.7% improvement is observed in protocol 1 as compared to other protocols in terms of throughput. Similarly, a minimum of 15.6% and maximum of 31.3% improvement is observed for protocol 1 as compared to other protocols.

References

1. Ashton, K.: That ‘Internet of Things’ thing, in the real world things matter more than ideas. *RFID J.* (2009). <http://www.rfidjournal.com/articles/view?4986>. Accessed 15 July 2014
2. Uckelman, D., Harrison, M., Michahelles, F.: *Architecturing the Internet of Things*. Springer, Berlin, Heidelberg (2011)
3. Aggarwal, C.C., Ashish, N., Sheth, A.: The Internet of Things: a survey from the data-centric perspective. In: Aggarwal, C. (ed.) *Managing and Mining Sensor Data*, pp. 383–428. Springer (2013)
4. Abyaneh, M.R.S.: Security analysis of lightweight schemes for RFID systems. Ph.D. thesis, University of Bergen, Norway (2012)

5. Juel, A., Weis, S.: Authenticating pervasive devices with human protocols. In: Shoup, V. (ed.) *Advances in Cryptology-Crypto 05*. LNCS, vol. 3126, pp. 293–298. Springer (2005)
6. Peris-Lopez, P., Hernandez-Castro, J.C., Estevez-Tapiador, J.M., Ribagorda, A.: RFID systems: a survey on security threats and proposed solutions. In: International Conference on Personal Wireless Communication-PWCA'06. LNCS, vol. 4217, pp. 159–170, Albacete, Spain (2006)
7. Moore, G.E.: Cramming more components onto integrated circuits. *Electronics* (1965). <http://www.intel.com>
8. Peris-Lopez, P., Hernandez-Castro, J.C., Estevez-Tapiador, J.M., Ribagorda, A.: Attacking RFID systems. In: Harold, F., Nozaki, K., Tipton, M. (ed.) *Information Security Management Handbook*, pp. 313–334. Auerbach Publications (2011)
9. Peris-Lopez, P., Hernandez-Castro, J.C., Estévez-Tapiador, J.M., Ribagorda, A.: LMAP: a real lightweight mutual authentication protocol for low-cost RFID tags. In: Proceedings of 2nd Workshop on RFID Security, p. 6 (2006)
10. Peris-Lopez, P., Hernandez-Castro, J.C., Estevez-Tapiador, J.M., Ribagorda, A.: M2AP: a minimalist mutual-authentication protocol for low-cost RFID tags. In: International Conference on Ubiquitous Intelligence and Computing, pp. 912–923. Springer, Berlin, Heidelberg (2006)
11. Li, T., Wang, G.: Security analysis of two ultra-lightweight RFID authentication protocols. In: IFIP International Information Security Conference, pp. 109–120. Springer, Boston, MA (2007)
12. Chien, H.Y.: SASI: a new ultralightweight RFID authentication protocol providing strong authentication and strong integrity. *IEEE Trans. Dependable Secure Comput.* **4**(4), 337–340 (2007)
13. Cao, T., Bertino, E., Lei, H.: Security analysis of the SASI protocol. *IEEE Trans. Dependable Secure Comput.* **6**(1), 73–77 (2009)
14. Phan, R.C.W.: Cryptanalysis of a new ultralightweight RFID authentication protocol—SASI. *IEEE Trans. Dependable Secure Comput.* **6**(4), 316–320 (2009)
15. Sun, H.M., Ting, W.C., Wang, K.H.: On the security of Chien's ultralightweight RFID authentication protocol. *IEEE Trans. Dependable Secure Comput.* **8**(2), 315–317 (2011)
16. Hernandez-Castro, J.C., Peris-Lopez, P., Phan, R.C.W., Tapiador, J.M.: Cryptanalysis of the David-Prasad RFID ultralightweight authentication protocol. In: International Workshop on Radio Frequency Identification: Security and Privacy Issues, pp. 22–34. Springer, Berlin, Heidelberg (2010)
17. Tian, Y., Chen, G., Li, J.: A new ultralightweight RFID authentication protocol with permutation. *IEEE Commun. Lett.* **16**(5), 702–705 (2012)
18. Cheng, Z.Y., Liu, Y., Chang, C.C., Chang, S.C.: Authenticated RFID security mechanism based on chaotic maps. *Secur. Commun. Netw.* **6**(2), 247–256 (2013)
19. D'Arco, P., De Santis, A.: On ultralightweight RFID authentication protocols. *IEEE Trans. Dependable Secure Comput.* **8**(4), 548–563 (2011)
20. Avoine, G.: Adversarial model for radio frequency identification. IACR Cryptol. ePrint Archive **2005**, 49 (2005)
21. Juels, A., Weis, S.A.: Defining strong privacy for RFID. *ACM Trans. Inf. Syst. Secur. (TISSEC)* **13**(1), 7 (2009)
22. Burmester, M., Van Le, T., de Medeiros, B.: Provably secure ubiquitous systems: universally composable RFID authentication protocols. In: Securecomm and Workshops, 2006, pp. 1–9. IEEE (2006)
23. Damgård, I., Pedersen, M.Ø.: RFID security: tradeoffs between security and efficiency. In: Topics in Cryptology—CT-RSA 2008, pp. 318–332. Springer, Berlin, Heidelberg (2008)
24. Ha, J., Moon, S., Zhou, J., Ha, J.: A new formal proof model for RFID location privacy. In: European Symposium on Research in Computer Security, pp. 267–281. Springer, Berlin, Heidelberg (2008)

Data Mining-Based Student's Performance Evaluator



Ravindra Kumar, Megha Kumar and Upasna Joshi

Abstract This paper is based on the study of different data mining techniques that are used to analyze and predict students' academic performance. The data mining application in education is the most promising tool for educational participation. The educational institution can apply this concept for students' performance analysis. In this study, the authors have collected students' academic data and three data mining algorithms related to classification, i.e., Naive Bayes, Decision Tree, and Convolutional Neural Network were used on the dataset. The prediction performance of three classifiers is compared and measured. This study will help educational institutes improve student academic performance.

Keywords Classification · Naive Bayes · Decision trees · Neural networks · CNN

1 Introduction

Current development in different domains has resulted in large amount of data collection that is usually stored in formats like sound, records, files, videos, and images. The collected data contributes in decision-making process, however analyzing/managing of large amount of huge data is very challenging. The data used in decision-making process requires proper process of knowledge extraction from large data repositories. To extract and discover relevant information from large amount of data, different data mining techniques can be used [1]. Different data mining algorithms and methods are used to mine patterns from data storage.

Wu [2] highlighted many data mining algorithms; using this, different algorithms effectiveness were analyzed by the authors and future research directions were outlined. Due to the significance of data mining in decision-making process, it can be successfully implemented in different domains like education [3, 4]. The purpose of this study is to discover different techniques to discover educational data that is further

R. Kumar · M. Kumar (✉) · U. Joshi

Department of Computer Science and Engineering, DTC, Greater Noida, India
e-mail: meghakumar245@gmail.com

used to determine the significance of learning systems [5], academic performance analysis [6], and development of warning system [7]. The analysis and academic performance prediction of student is important for student academic progress [3, 8] and proves to be a difficult task because of the control of several factors like psychological profile, family factor, and student interaction with their teachers and classmates [9].

The current methods of data mining can be generally categorized into five categories, out of which one is predication that deals with output prediction based on input parameters. Predication can be divided into three categories: (1) Density estimation, (2) Regression, and (3) Classification. Some of the popular classification algorithms includes Naive Bayes, Neural Networks, Support Vector Machines, and Decision Trees with their predication is either categorical or binary variable [10]. In this paper, the authors have used classification algorithms, i.e., Decision Tree, Naive Bayes, and Convolutional Neural Network for predicting students' academic performance for undergraduate courses. The authors have collected students' academic data from Computer Science Department of their college and further, the performance prediction results of three classifiers will be analyzed and compared.

The paper is organized in five sections, which is described as follows. The first part outlines the importance of data mining in the field of education. Sections 2–4 discuss the literature review of different data mining algorithms. At the end, conclusion is drawn and future research is presented.

2 Literature Review

There are various techniques and algorithms that are used for extraction of knowledge from databases related to education. Some of the algorithms are described as follows:

1. Classification

Classification is the most common technique used. In this technique, a model is developed to assign class or label to a record from pre-classified examples. The technique is divided into two parts: training and testing part. In training, a model is developed with a portion of the dataset called as training set that knows the attributes and classes. When the model is built, it is further used to classify a new record to a class when the class is not known.

Some of the algorithms to design a classifier model includes Naive Bayes, Decision Tree (DT), Support Vector Machines, and Neural Network (NN) [11]. The authors have discussed Convolutional Neural Networks, Decision Tree and Naïve Bayes in this study.

2. Decision Trees

A decision tree is a flowchart-like tree structure, where the leaf nodes are represented by ovals and internal nodes are represented using rectangles. All the internal nodes have two or more child nodes and contain branches, where the value of an expression

of the attributes is tested. The distinct outcomes of the test are used to label the arcs from an internal node to its children. Each leaf node is associated with a class label.

The data tuples of the training set are used to develop a decision tree. Each tuple of the training set is represented by attributes set and a class label. The attributes can hold values that are continuous or discrete. Decision tree classifies the data tuples to the classes whose classes are not known and the path from the root to the leaf is followed based on the attribute values of the tuple. The class predicted by the decision tree is the class named to that tuple. Some of the decision trees building algorithms are: C4.5, ID3, and CART.

2.1 Classification Using Decision Tree Induction

The procedure of construction of a tree from the training set is called as tree induction. A greedy (non backtracking) top-down divide and conquer method is adopted by many existing tree induction systems. Starting with the entire training set and an empty tree, the following algorithm is implemented on the training data, until no more splits are possible [12, 13].

2.2 Decision Tree Algorithms

ID3, a greedy algorithm was introduced by Quinlan [14] where the information gain is used in the selection of next attributes. The attribute holding greatest entropy or highest information gain is selected as the current node test attribute.

C4.5, a successor of ID3 is the most popular decision tree algorithm. It is an improvement over ID3. C4.5 handles both continuous and discrete attribute and uses Gain ratio [15] as an attribute selection measure.

CART algorithm, proposed by Breiman, is similar to ID3. The Gini index [14] an impurity measure is used in selection of attribute in CART. For the nominal target variable, it generates a classification tree and a regression tree is constructed for continuous-valued numerical target variable.

The difference lies in split criterion among three algorithms and the well-known split criteria are Information Gain, Gain Ratio, and Gini Index.

The split criteria used in ID3 is Information Gain, which is given in Eq. (1). When an attribute A splits the set S into subsets Si.

$$Gain(S, A) = E(S) - I(S, A) = E(S) - \sum \frac{|S_i|}{|S|} \cdot E(S_i) \quad (1)$$

Gain Ratio, an extension of information gain which decreases its bias towards multi-valued attributes is represented in Eq. (2):

$$GainRatio(S, A) = \frac{Gain(S, A)}{GainIf(S, A)} \quad (2)$$

Classification and Regression Trees (CART) algorithm uses the Gini Index as mentioned in Eq. (3) to find impurities.

$$Gini(S) = 1 - \sum_i p_i^2 \quad (3)$$

3 Artificial Neural Networks with Deep Learning

AI emerges as a new technology with various subfields as deep learning, machine learning, etc. Due to advanced technology and various applications in every field of life, it proved to be advantageous for everyone. The various real times application like handwriting reorganization, understating the features of images, understanding the traffic on roads, for example GPS, object detection for security and military purposes. AI techniques, approaches and complex algorithms make the complex problems easier and increase the accuracy to solve the problems. [16].

Many approaches are there in the field of AI. But basically, two major classifications can be done in this area. The first approach includes reasoning based upon logic and logic programming. The second one includes biological models artificially designed in AI to solve the problems [17]. The genetic algorithm and artificial neural networks (ANNs) are considered to be the major approaches implemented under the second type of category. In ANN, neural nodes are trained to gain the assigned weight using the learning it stores the experimental knowledge and makes it available to use [18, 19]. Neural networks can be extended to various layers input layer, many hidden layers, and output layers. The grid is trained to find the relationships between the adoptable nodes.

Convolutional neural networks also known as CNNs, have one or several hidden layers compared to simple neural networks. The network features include highly varying and nonlinear functions These types of neural networks are comprised of many connections, and the architecture includes various types of layers like pooling and different layers of network [21]. Deep neural networks having numerous neural nodes and comprised of layerwise nonlinear mathematical calculations. In [22], every CNN architecture represents different level of abstraction features.

Convolutional Neural Networks have marked tremendous success in the areas like sentence and text classification [23], image recognition [24, 25], data analytics and time series [26, 28], medicine [27], error diagnosis [29] and biological image classification [31]. Many companies like Apple, Adobe, Facebook, IBM, Google, Baidu, Microsoft, Netflix, NEC, and NVIDIA [32] have used deep learning techniques.

4 Naïve Bayes

Naïve Bayes: A statistical and probabilistic model based upon the Bayes Theorem [33] with the independent features. The Naïve Bayes is generally applicable on large data set, which has a very fast response and performance predictable results as compared to other models. The evaluation of probability model depends upon the posterior probability and prior probability. To calculate the posterior probability, the equation for this theorem is stated below

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} \quad (4)$$

Let x be a data tuple. $P(y/x)$ is a posterior probability, of y conditioned on x . In contrast, $P(x)$ is the prior probability, or a priori probability, of x . Similarly, $P(x/y)$ is the posterior probability of x conditioned on y . $P(y)$ is the prior probability of x . Bayes' theorem is useful in that it provides a way of calculating the posterior probability, $P(y/x)$, from $P(y)$, $P(x/y)$, and $P(x)$.

Naïve Bayes classifier is useful for many applications either it can be industry, military, or medical field. Due to its simplicity and robustness, this algorithm is widely used for fault prediction, image classification, text or document classification, fraud detection, etc.

Text classification is an important part of natural language processing. In this paper [34], Naïve Bayes used for text classification is used with Gaussian, Bernoulli, and classical naïve Bayes event model. The comparison is done between these three models as results evaluate Gaussian event model performed better than other two models.

Arar et al. [35] proposed the features-based Naïve Bayes method for classification to detect the defect in the software. For this, experiments were conducted on the NASA Promise dataset. The results shows that current model is comparatively better than the classical Naive Bayes classifier.

In case of gene expressions [36], the dataset is contaminated by outliers, thus in this paper, the gaussian NB method is used for robustify by β -divergence method. The aim of this method is to produce the estimate scale and location parameters which are based upon the outliers and modification in test data. The comparison done between classical NB, KNN, SVM, and Adaboost proved that proposed method is better than others.

Al-khurayji et al. [37] defines a method based upon Kernal Naïve Bayes (KNB) classifier for text classification of Arabic nonlinear problem. The comparison is done between different classifiers but this method shows the effectiveness than other models.

In biomedical study, the extraction of entity (gene, disorder, virus) is an important part of the study. For this approach, machine learning algorithms are helpful and extraction the features. The chemical compound features are extracted in [38], using

Naïve Bayes and wrapper method. The results show that the combination of both the methods is helpful identifying the best features and comparative results are much better.

5 Conclusion

Data Mining and machine learning-based techniques like Decision Tree, Artificial Neural Networks, and Naive Bayes are useful for predication and analyzing the student's performance based upon academic outcomes. This study is applicable for faculty to predict and monitor the students who are about to fail in the upcoming examinations. Different studies have been conducted to identify different factors like family, personal, and instructor factor that affect students' performance. The factors can vary from one institute to another, one country to another, and one culture to another. Amongst all, the instructors' role is very important as they provide proper guidance and motivation to the students. For the future scope, the study will be including the large volume of dataset from different courses and academic fields for investigations.

Acknowledgements The authors would like to acknowledge the Department of Computer Science and Engineering, DTC, Greater Noida, India for providing student's academic data and for the necessary permissions to use this in our research studies.

References

1. Essa, A., Ayad, H.: Student success system: risk analytics and data visualization using ensembles of predictive models. In: Paper Presented at the 2nd International Conference on Learning Analytics and Knowledge, Vancouver (2012)
2. Kindong, W., Vipin Kumar, J., Quinlan, R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G.J., Ng, A.F.M., Liu, B., Yu, P.S., Zhou, Z.-H., Steinbach, M., Hand, D.J., Steinberg, D.: Top 10 algorithms in data mining. *Knowl. Inf. Syst.* **14**(1), 1–37 (2008)
3. Baker, R.S., Corbett, A.T., Koedinger, K.R.: Detecting student misuse of intelligent tutoring systems. In: Proceedings of the 7th International Conference on Intelligent Tutoring Systems, pp. 531–540 (2004)
4. Zaïane, O.: Building a recommender agent for e-learning systems. In: Proceedings of the International Conference on Computers in Education, pp. 55–59 (2002)
5. Kotsiantis, S.B.: Use of machine learning techniques for educational proposes: a decision support system for forecasting students grades. *Artif. Intell. Rev.* **37**(4), 331–344 (2012)
6. Hongbo, D., Yizhou, S., Yi, C., Jiawei, H.: Probabilistic models for classification. In: Data Classification, pp. 65–86. Chapman and Hall/CRC (2014)
7. Romero, C., Ventura, S.: Data mining in education. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **3**(1), 12–27 (2013)
8. Araque, F., Roldan, C., Salguero, A.: Factors influencing university dropout rates. *J. Comput. Educ.* **53**, 563–574 (2009)
9. Baker, R.S.J.D.: Data mining for education. In: McGaw, B., Peterson, P., Baker, E. (eds.) *International Encyclopedia of Education*, 3rd edn., vol. 7, pp. 112–118 (2010)

10. Manzoor, U., Nefti, S.: An agent based system for activity monitoring on network-ABSAMN. *Expert Syst. Appl.* **36**(8), 10987–10994 (2009)
11. Alkhasawneh, R., Hobson, R.: Modeling student retention in science and engineering disciplines using neural networks. In: Proceedings of the IEEE Global Engineering Education Conference (EDUCON), pp. 660–663 (2011)
12. Han, J., Kamber, M.: Data Mining Concepts and Techniques, 2nd edn. Morgan Kaufmann, SanFrancisco. ISBN: 978-81-312
13. Quinlan, J.R.: Induction of decision trees. In: *Machine Learning*, vol. 1, pp. 81–106. Morgan Kaufmann (1986)
14. Kohavi, R., Quinlan, R.: Decision tree discovery. In: *Handbook of Data Mining and Knowledge Discovery*. University Press (1999)
15. Soman, K.P., Diwakar, S., Ajay, V.: *Insight into Data Mining-Theory and Practice*. Prentice Hall of India, New Delhi. ISBN: 81-203-2897-3
16. Goodfellow, I., Bengio, Y., Courville, A.: *Deep Learning (Adaptive Competition and Machine Learning)*, p. 779. The MIT Press (2016)
17. Munakata, T.: *Fundamentals of the New Artificial Intelligence: Neural, Evolutionary, Fuzzy and More*, 2nd edn., p. 225. Springer, London (2008)
18. Floreano, D., Dürre, P., Mattiussi, C.: Neuroevolution: from architectures to learning. *Evol. Intel.* **1**(1), 47–62 (2008). <https://doi.org/10.1007/s12065-007-0002-4>
19. Prieto, A., Atencia, M., Sandoval, F.: Advances in artificial neural networks and machine learning. *Neurocomputing* **121**, 1–4 (2013). <https://doi.org/10.1016/j.neucom.2013.01.008>
20. Dalto, M.: Deep Neural Networks for Time Series Prediction with Application in Ultra-Short-Term Wind Forecasting, pp. 1657–1663. IEEE (2015)
21. Ferreira, A., Giraldo, G.: Convolutional neural network approaches to granite tiles classification. *Expert Syst. Appl.* **84**, 1–11 (2017). <https://doi.org/10.1016/j.eswa.2017.04.053>
22. Bengio, Y.: Learning deep architectures for AI. *Found. Trends Mach. Learn.* **2**(1), 1–127 (2009). <https://doi.org/10.1561/2200000006>
23. Conneau, A., Schwenk, H., Barrault, L., Lecun, Y.: Very deep convolutional networks for text classification. In: Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, vol. 1, Long Papers (2017). <https://doi.org/10.18653/v1/e17-1104>
24. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: Proceedings of Neural Networks (NIPS), Nevada, USA, pp. 1106–1114 (2012)
25. Simonyan, K., Zisserman, A.: Very Deep Convolutional Networks for Large-scale Image Recognition. Published as a conference paper at ICLR. Cornell University Library (2015)
26. Zheng, Y., Liu, Q., Chen, E., Ge, Y., Zhao, J.L.: Time series classification using multi-channels deep convolutional neural networks. In: *Lecture Notes in Computer Science*, pp. 298–310 (2014). https://doi.org/10.1007/978-3-319-08010-9_33
27. Acharya, U.R., Fujita, H., Oh, S.L., Hagiwara, Y., Tan, J.H., Adam, M.: Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals. *Inf. Sci.* **415–416**, 190–198 (2017). <https://doi.org/10.1016/j.ins.2017.06.027>
28. Wang, K., Zhao, Y., Xiong, Q., Fan, M., Sun, G., Ma, L., Liu, T.: Research on healthy anomaly detection model based on deep learning from multiple time-series physiological signals. *Sci. Program.* **2016**, 1–9 (2016). <https://doi.org/10.1155/2016/5642856>
29. Liu, R., Meng, G., Yang, B., Sun, C., Chen, X.: Dislocated time series convolutional neural architecture: an intelligent fault diagnosis approach for electric machine. *IEEE Trans. Industr. Inf.* **13**(3), 1310–1320 (2017). <https://doi.org/10.1109/iti.2016.2645238>
30. Meng, M., Chua, Y.J., Wouterson, E., Ong, C.P.K.: Ultrasonic signal classification and imaging system for composite materials via deep convolutional neural networks. *Neurocomputing* **257**, 128–135 (2017). <https://doi.org/10.1016/j.neucom.2016.11.066>
31. Affonso, C., Rossi, A.L.D., Vieira, F.H.A., de Leon Ferreira, A.C.P.: Deep learning for biological image classification. *Expert Syst. Appl.* **85**, 114–122 (2017). <https://doi.org/10.1016/j.eswa.2017.05.039>

32. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darell, T.: Caffe: Convolutional Architecture for Fast Feature Embedding. Cornell University Library (2014)
33. Rish, I.: An empirical study of the naive Bayes classifier. In: IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, vol. 3. no. 22. IBM, New York (2001)
34. Xu, S.: Bayesian Naïve Bayes classifiers to text classification. *J. Inf. Sci.* **44**(1), 48–59 (2018)
35. Arar, Ö.F., Ayan, K.: A feature dependent Naive Bayes approach and its application to the software defect prediction problem. *Appl. Soft Comput.* **59**, 197–209 (2017)
36. Ahmed, Md., et al.: Robustification of Naïve Bayes classifier and its application for microarray gene expression data analysis. *BioMed Res. Int.* **2017** (2017)
37. Al-khurayji, R., Sameh, A.: An effective Arabic text classification approach based on Kernel Naïve Bayes classifier. *Int. J. Artif. Intell. Appl.* 01–10 (2017)
38. Alshaikhdeeb, B., Ahmad, K.: Feature selection for chemical compound extraction using wrapper approach with Naive Bayes classifier. In: 2017 6th International Conference on Electrical Engineering and Informatics (ICEEI). IEEE (2017)

A Comprehensive Review on the Issues Related to the Data Security of Internet of Things (IoT) Devices



**Shivansh Upadhyay, Shashwat Kumar, Sagnik Dutta,
Ajay Kumar Srivastava, Amit Kumar Mondal and Vivek Kaundal**

Abstract Internet of Things or IoT is an emerging Internet-based structure of various devices that allows exchange, processing and storage of data along with various other applications. With IoT dominating the industrial domains around the world, followed by its increased implementation for completing various task. Furthermore, it is also prone to malicious attacks that show the capability of the IoT devices in terms of security and privacy of the data. Though previous research in this area may have explored and brought forth the security challenges faced by IoT but there are still some gaps remain in a few places. Therefore, it is very important to design such a framework that is having the capability to face the challenges related to security. In this paper, a comprehensive analysis of security threats faced by IoT and the possible solutions are suggested to solve the issues. We ensure to conduct proper research on the legal challenges of IoT. We believe our approach can further help others to solve the threats easily.

Keywords Internet of Things · Data security · Communication protocol · Wireless networks

S. Upadhyay · S. Kumar · S. Dutta · A. K. Mondal (✉) · V. Kaundal
Department of Electrical and Electronics Engineering,
University of Petroleum and Energy Studies, Dehradun, India
e-mail: akmondal1603@gmail.com

S. Upadhyay
e-mail: shivansh1018@gmail.com

S. Kumar
e-mail: shashwatray773@gmail.com

S. Dutta
e-mail: sagnikdutta134@gmail.com

V. Kaundal
e-mail: vkaundal@ddn.upes.ac.in

A. K. Srivastava
Department of Mechanical Engineering,
University of Petroleum and Energy Studies, Dehradun, India
e-mail: akumar@ddn.upes.ac.in

1 Introduction

Taking a glimpse at the modern-day scenario of newly emerging technology, almost every electronic device is connected to the internet, or has the potential to do so. Apart from increasing the functionality and utility of regular electronic devices, this also exposes them to the open-ended side of the vast internet, in which they are merely a node [1, 2]. This makes them vulnerable to various attacks and breaches. This further gives rise for a need to create a robust security measure for these devices.

Getting connected to the internet follows certain protocols, the basic of which is obtaining an IP address within the network. This grants it access to the internet through computers present locally in the network, which generally act as gateways for it. Generally, electronic IOT devices lack resources such as RAM and clock frequency to perform complex calculations and algorithms required for encryptions in device-to-device communications. Therefore, this makes it easy for hackers to passively sniff and intercept these communications, steal confidential data and sometimes even alter the data, which can cause various degrees of losses to the user, as well as to the cloud service provider [3–5].

In many cases, users often do not change the default passwords of these systems or set easily guessable passwords. This can cause a simple algorithm to grant unauthorized access to hackers. User errors make such devices vulnerable to such attacks, even if secure communication techniques are established. The best example is the Mirai botnet, which hit the internet like a storm in 2016. It was programmed to run a sweep across the internet for such devices with open ports, checking from a well-defined set of most commonly set passwords using telnet protocol [6]. On successful match and access gain, it spread the virus to devices connected with the originally compromised device. It is predicted that by 2050 as shown in Fig. 1, all devices will have the ability to perform sensor-based communication, which will increase the methods of penetrating into their security structures, which consecutively increases the need to implement strong security measures [7].

2 Methodology and Discussion

2.1 Security Constraints and Requirements

2.1.1 Security Constraints

IoT devices are intrinsically resource constrained. In this way, utilizing the standard security parts particularly in the brilliant things isn't direct. The significant security requirements of IoT gadgets are clearly depicted in Fig. 2 and is explained as follows:

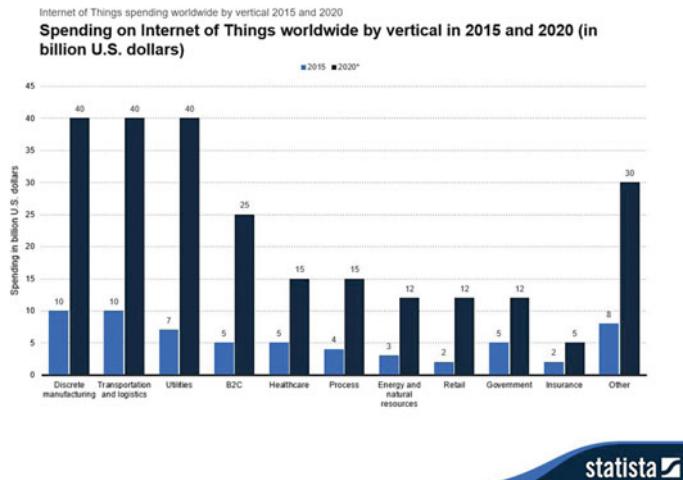
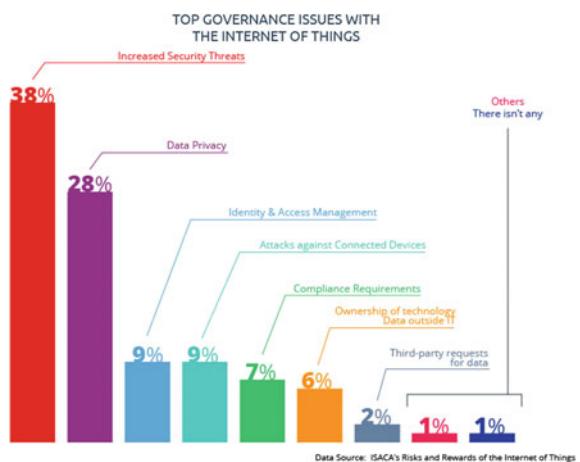


Fig. 1 Worldwide implementation of IoT devices (2015–2020) [8]

Fig. 2 Technical issues related to Internet of Things (IoT) based devices



2.1.2 Constraints Dependent on Equipment

- (i) **Computational and vitality requirement:** Most of the time, in IoT gadgets which are battery powered and devices nowadays, low-control CPUs with low clock rate are being used. Thus, computationally expensive cryptographic calculations requiring quick computation cannot be ported specifically to these low fuelled devices [9].
- (ii) **Memory requirement:** IoT devices are worked with obliged Smash and Flash memory appeared differently in relation to the standard propelled structure (e.g. Laptop, PC and so on.), and use Real-Time Operating System (RTOS)

or lightweight type of universally useful framework. They too run structural programming and restrictive organizations. Thus, security designs should be memory productive. Nevertheless, standard security counts are not planned specifically thinking about the memory capability, claiming the customary computerized framework utilizes spacious RAM and hard drive. For those security designs, enough memory space will not be available in the wake of booting up the working structure and system programming. In this manner, regular security computations cannot be used specifically to anchor IoT gadgets [10].

- (iii) **Resistance from tempering of devices:** Many times, devices are installed in remote areas and abandoned. Here, attackers can tamper them and extract the data and alter them accordingly, and Temper resistance package can be installed to safeguard them from the attackers.

2.2 Security Requirement

There are several security requirements enlisted below.

2.2.1 Data Security Necessities

- (i) **Integrity:** An enemy can change the information and bargain the trustworthiness of an IoT framework. Therefore, honesty ensures that any obtained data hasn't been modified in travel [11].
- (ii) **Information security:** Secrecy and privacy of on-air and put away information should be entirely spared. It alludes to restricting the information get to and exposing to the approved IoT hub, and averting access by unauthorized ones or exposure to them. For example, an IoT system should not reveal the readings of its sensors to its neighbours (if it is arranged not to do as such).
- (iii) **Anonymity:** Anonymity covers the source of the data. This security benefit helps with the data protection and security.
- (iv) **Non-renouncement:** Non-disavowal is the affirmation that somebody cannot deny something. For example an IoT hub can't deny communicating something in particular that it has already sent.
- (v) **Freshness:** It is needed to ensure each message's freshness. Freshness guarantees that the information is in unique, and it has not replayed any old messages.

2.2.2 Access Level Security Prerequisites

- (i) **Authentication:** Authentication empowers an IoT device to ensure the nature of the associate with which it conveys (e.g. recipient checks that the got information began from the rightful source or not). It additionally needs guaranteeing that legitimate clients gain admittance to the IoT devices and the systems for

managerial undertakings: remote reconstructing or on the other hand to control the IoT devices and systems [12].

- (ii) **Authorization:** It ensures that just the approved gadgets furthermore, the clients gain admittance to the system administrations or assets.
- (iii) **Access control:** Access control is the demonstration of guaranteeing that a confirmed IoT hub gets to just what it is approved to, and that's it.

2.3 Utilitarian Security Prerequisites

- (i) **Exception taking care of:** Exception dealing with affirms that an IoT organize is alive and keeps serving even in the odd circumstances: hub trade-off, hub obliteration, failing equipment, programming glitches, separation ecological perils and so on. Consequently, it guarantees heartiness.
- (ii) **Availability:** Availability guarantees the survivability of IoT administrations to approved gatherings when required regardless of denial-of-benefit assaults. It likewise guarantees that it has the capacity to give a base level of administrations in the nearness of intensity misfortune, disappointments [13].
- (iii) **Resiliency:** On the off chance that a couple of bury associated IoT gadgets are endangered, a security plan should, in any case, ensure against the assault.
- (iv) **Self-association:** An IoT gadget may fall flat or run out of vitality. The rest of the gadget or colleague gadgets ought to be able to be redesigned to keep up a set level of security.

2.4 Attack Strategy and Information Damage Level, with Possible Solutions

Various attacking groups use different strategies for device penetration, which operate varied interests of the attacker.

The attacks based on different levels of information disruption are:

- (i) **Man-in-the-middle attacks:** These are passive attacks which allow the attackers to act as 'middlemen' in the communication passage from the IOT device to the commuter system. The hacker uses vulnerabilities in Address resolution protocol, causing the device to interpret the hacker as the commuter, and vice versa for the commuter. Proper packet administration is required to prevent such attacks [14]
- (ii) **Alteration:** When performing these attacks, the attacker changes information in the database, which can cause confusion in interpreting the protocol, and compromises the integrity of data. Fail proof network management tools are necessary to prevent such attacks.

- (iii) **Message Replay attacks:** These attacks confuse the protocol interpreters by intercepting and importing messages out-of-the-protocol timing. Efficient systems to eliminate message replay and traceability and cloning attacks are needed to prevent such attacks.
- (iv) **Interruption:** Basically, these are common Denial-of-Service attacks such as shutting out power sources, blocking data transfer. The only way to prevent such attacks is implementing efficient disaster recovery methods.

2.4.1 Host Based Attacks

These attacks are directly pointed towards the working machinery system governing the actual functioning of the device [15]. These are of three types:

- i. **User compromise:** The attacker directly withdraws important and confidential security credentials by entrapping those using forceful means. Securely storing important credentials is needed to manage the aftermaths of such attacks.
- ii. **Hardware compromise:** Important credentials and information is extracted directly from the IOT device by successfully distorting its hardware accordingly. A tamper-resistant design in the internal architecture of IOT devices can ensure the prevention of such attacks.
- iii. **Software compromise:** Directly accessing open ports in the device's operating system, can cause loss of data integrity, and can also give attackers a chance to damage the kernel. Implementing secure shell connections and introducing strong processing architecture to apply cryptographic methods can prevent such compromises.
- iv. **Fabrication:** A fabrication attack involves creating and inserting false and illegitimate information along with data and processes within the system. These false data are inserted alongside the authentic data thus creating false record and similar other things. Taking regular accurate survey of the authentic data can help prevent fabrication.
- v. **Protocol disruption:** The attacker might act from inside or outside the network and perform illegal activities on standard conventions: key administration convention, information accumulation convention, synchronization convention, and so forth.
- vi. **Eavesdropping:** The attacker can gain access to any information via private correspondence once inside the network. Radio Frequency Identification (RFID) devices are the most standouts among the IoT devices to be attacked.
- vii. **Physical Attacks:** These types of attacks involve causing physical damage to the device thus altering the IoT device and causing vindictive code infusion. Maximizing security around devices can prevent these attacks.

3 Conclusion

A comprehensive review on the most vital security parts of the Internet of Things (IoT) with an accentuation on what is being done and what are the issues that require further examination. Our work investigates the general security design of IoT pursued by security issues identified with interoperability of heterogeneous items. We likewise play out a comprehensive investigation of the vulnerabilities of the associated questions by taking thought of their computational constraint, vitality constraint, asset confinement, and lightweight cryptographic conventions. We have addressed genuine circumstances where the absence of IoT security could present different dangers. Our work breaks down existing exploration issues and challenges and gives chances to future research work around there. In end, we trust this overview may give a vital commitment to the examination network, by recording the ebb and flow security status of this extremely unique territory of research also, spurring scientists inspired by growing new plans to address security with regards to the Internet of Things.

References

1. Gubbi, J., et al.: Internet of Things (IoT): a vision, architectural elements, and future directions. *Futur. Gener. Comput. Syst.* **29**(7), 1645–1660 (2013)
2. Dehury, C.K., Sahoo, P.K.: Design and implementation of a novel service management framework for IoT devices in cloud. *J. Syst. Softw.* **119**, 149–161 (2016)
3. Santoso, F.K., Vun, N.C.: Securing IoT for smart home system. In: 2015 IEEE International Symposium on Consumer Electronics (ISCE). IEEE (2015)
4. Sivaraman, V., et al.: Network-level security and privacy control for smart-home IoT devices. In: 2015 IEEE 11th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob). IEEE (2015)
5. Atzori, L., Iera, A., Morabito, G.: The Internet of Things: a survey. *Comput. Netw.* **54**(15), 2787–2805 (2010)
6. Borgohain, T., Kumar, U., Sanyal, S.: Survey of security and privacy issues of Internet of Things (2015). [arXiv:1501.02211](https://arxiv.org/abs/1501.02211)
7. Solà Campillo, O.: Security Issues in Internet of Things (2017)
8. Kelly, S.D.T., Suryadevara, N.K., Mukhopadhyay, S.C.: Towards the implementation of IoT for environmental condition monitoring in homes. *IEEE Sens. J.* **13**(10), 3846–3853 (2013)
9. Sicari, S., et al.: Security, privacy and trust in Internet of Things: the road ahead. *Comput. Netw.* **76**, 146–164 (2015)
10. Sehgal, A., et al.: Management of resource constrained devices in the Internet of Things. *IEEE Commun. Mag.* **50**(12) (2012)
11. Suo, H., et al.: Security in the Internet of Things: a review. In: 2012 International Conference on Computer Science and Electronics Engineering (ICCSEE). IEEE (2012)
12. Baccelli, E., et al.: RIOT OS: towards an OS for the Internet of Things. In: 2013 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). IEEE (2013)

13. Raza, S.: Lightweight Security Solutions for the Internet of Things. Mälardalen University, Västerås, Sweden (2013)
14. Weber, R.H.: Internet of Things—new security and privacy challenges. *Comput. Law Secur. Rev.* **26**(1), 23–30 (2010)
15. Covington, M.J., Carskadden, R.: Threat implications of the Internet of Things. In: 2013 5th International Conference on Cyber Conflict (CyCon). IEEE (2013)

SVD-Based Linear Precoding Using Channel Estimation for MIMO-OFDM Systems



R. Raja Kumar, R. Pandian, C. Satheeswaran, M. KaviPriya and P. Indumathi

Abstract Orthogonal Frequency Division Multiplexing is an approach that can be used to combat the effect of frequency selective fading channels. Apart from spectral efficiency, reliability over fast fading channels is a primary concern in OFDM. MIMO systems offer a solution to this problem owing to the high capacity and beamforming gain. The integration of MIMO with the traditional OFDM systems allow parallel transmission of data with a significant reduction in inter-carrier interference. The tradeoff to be considered is the increase in complexity of the user equipment. This is due to the advanced modulation and estimation operations performed at the receiver. In this paper, we present a transmitter site channel estimation scenario with the help of pilot symbols. Further, a MIMO-OFDM transceiver is designed which utilizes the information of the channel state at the transmitter to perform Singular Value Decomposition (SVD) based linear precoding on the data. Such a type of precoding scheme provides reduction in bit error rate to an optimum range. The simulation results show that a singular value-based precoding technique performs better than the traditional Alamouti space–time block code. But the requirement of additional time slot for channel estimation is a challenge that accompanies this technique.

Keywords MIMO-OFDM · SVD · Alamouti coding · QR precoding · MMSE

R. Raja Kumar (✉)

Mathematics Department, Sathyabama Institute of Science and Technology, Chennai, India
e-mail: rkmird@gmail.com

R. Pandian

Department of Electronics and Instrumentation Engineering, Sathyabama Institute of Science and Technology, Chennai, India

C. Satheeswaran

Department of Electronics and Communication Engineering, Dhaanish Ahmed College of Engineering, Chennai, India

M. KaviPriya · P. Indumathi

Department of Electronics Engineering, Anna University, MIT Campus, Chennai, India

1 Introduction

Current generation wireless systems heavily rely on Multiple-In-Multiple Out (MIMO) technology for obtaining higher data rate and increased SNR for data transmission. As the number of antennas at the receiver side increases, the diversity order increases proportionally. There has been considerable attention given to the capacity and performance of MIMO systems with varying degrees of CSI available at the transmitter. However, less attention has been given to the methods used for obtaining this information at the transmitter. Here, Thiagarajan proposed a novel transmit precoding technique in order to increase the diversity.

In this paper we are designing a MIMO Orthogonal Frequency Division Multiplexing (OFDM) transceiver and perform channel estimation at the transmitter site and analyze the performance for different modulation schemes like BPSK, QPSK, 16-QAM using precoding techniques like SVD, Alamouti coding and QR-based schemes. Also, the plot for capacity curve and calculation of the time complexity of various estimation techniques are done and analyzed.

2 Related Work

The idea of basic wireless communications is explained by Rappaport in [7]. The basics of multiple element antennas (MEA) is explained by Molisch in [6]. Jose et al. [4] suggested the idea of exploiting the reciprocity of the channel in TDD systems. Thiagarajan et al. [9] provided that infinite diversity order can be achieved when the CSI (Channel State Information) is available only at the transmitter. A novel transmit precoding based on QR decomposition with number of transmit antennas being twice as that of receive antennas is developed. The numerical results illustrate that the CSIT-based transmission schemes offer better performance in terms of BER as compared to CSIR-based schemes.

This paper consists of the literature existing for the possibility of the transmitter site estimation. Further, an implementation of the system using precoding to further decrease the error rate highlights the necessity to move towards such a system.

3 System Model

The model considered here is a 2×2 MIMO-OFDM system with channel estimation at the transmitter side as described in Fig. 1.

The channel is estimated using the pilot symbol transmission. Binary Phase-Shift Keying Modulation is used for ubiquitous transmission of pilot symbols. This is because BPSK allows higher accuracy required for exact estimation of channel properties. An adaptation function f is designed which adapts to the channel h . An

Fig. 1 A simplified 2×2 MIMO system

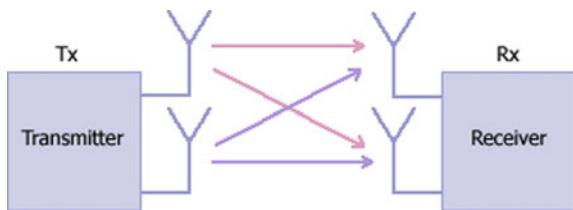


Fig. 2 Adaptation function

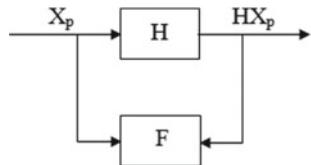
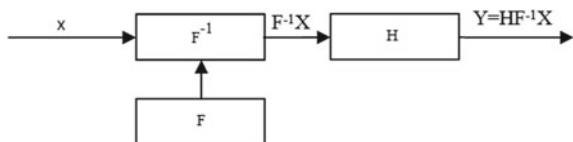


Fig. 3 Data symbols transmission



inverse filter F^{-1} has a frequency response of F^{-1} . Thus, each of the incoming pilot is multiplied by this factor, and then transmitted to the receiver. In essence, this pre-distortion type adaptive equalization nullifies the effect of the channel as illustrated in Figs. 2 and 3.

Some of the assumptions taken into account are that channel is a mirror channel. This means that channel properties like path loss and fading are same for both forward and the reverse channel. Also, the channel is assumed to be slow fading. Hence, the response of the channel is constant for a given coherence time.

4 Estimation Algorithms and Precoding

4.1 Zero Forcing

The zero forcing-based estimation algorithm is simplest but provides the least performance when compared to other techniques. First, the pilot bits are transmitted from the receiver to transmitter. Since this information is known at both sides, the received signal can be considered to have similar properties of the channel, thus providing the instantaneous channel response. The mathematical representation can be provided as

$$y = Hx + n$$

where H is the channel response and y is the received signal. To solve this equation, we estimate a matrix M , which is the pseudoinverse matrix such that

$$M * H = I$$

This constraint can be met by the following linear estimator:

$$\begin{pmatrix} h_{1,1}^* & h_{2,1}^* \\ h_{1,2}^* & h_{2,2}^* \end{pmatrix} \begin{pmatrix} h_{1,1} & h_{1,2} \\ h_{2,1} & h_{2,2} \end{pmatrix} = \begin{pmatrix} |h_{1,1}|^2 + |h_{2,1}|^2 & h_{1,1}^* h_{1,2} + h_{2,1}^* h_{2,2} \\ h_{1,2}^* h_{1,1} + h_{2,1} h_{2,2}^* & |h_{1,2}|^2 + |h_{2,2}|^2 \end{pmatrix}$$

4.2 Minimum Mean Square Estimation (MMSE)

This type of estimator is similar to zero forcing but an additional term for noise variance is added. As a result, the bit error rate reduces to an extent. In this estimation,

$$M = (H^H H + N_0 I)^{-1} H^H \quad (1)$$

where H^H is the hermitian transpose of the channel matrix.

4.3 Maximum Likelihood Estimation

The ML-based channel estimation produces the best performance in terms of error. But, for higher order modulation techniques, a larger codebook design is required. Also, the comparison of the data for each symbol detection provides an increased level of system complexity.

$$J = \left| \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \begin{pmatrix} h_{1,1} & h_{1,2} \\ h_{2,1} & h_{2,2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right|^2$$

ML scheme computes J for every possible combination of symbol values. These element values have to be stored and the minimum value among these gives an accurate estimate of the received signal.

4.4 QR-Based Estimation

In this system, the channel matrix is decomposed to QR form and based on the least square approximate method the desired signal transmitted is estimated. The detected desired signal x_{qlqr} is computed as

$$\begin{aligned}
 X_q &= ((R^H R + \sigma^2 I)^{-1} H^H) * y \\
 r &= y - H * x_q \\
 e &= ((R^H R + \sigma^2 I)^{-1} H^H) * r \\
 x_{q1qr} &= x_q + e
 \end{aligned} \tag{2}$$

This system provides a better performance compared to other systems except ML. However, the simplicity in terms of complexity and design makes this an attractive option for MIMO-based data transfer techniques.

5 Precoding

5.1 Singular Value-Based Precoding

Once the channel from the receiver to transmitter is estimated, assuming that it is a reciprocal channel, singular value decomposition is performed on the channel values. Thus, matrices U, Σ and V are obtained. Here, Σ is the diagonal matrix which converts the Rayleigh channel to parallel flat fading channel matrix thus increasing the diversity.

Figure 4 illustrates the system model considered. Here, a 64-point FFT is performed in each of the transmit antenna. The matrix U has to be known at the receiver. This results in a separate transmission of this matrix and detection at the receiver.

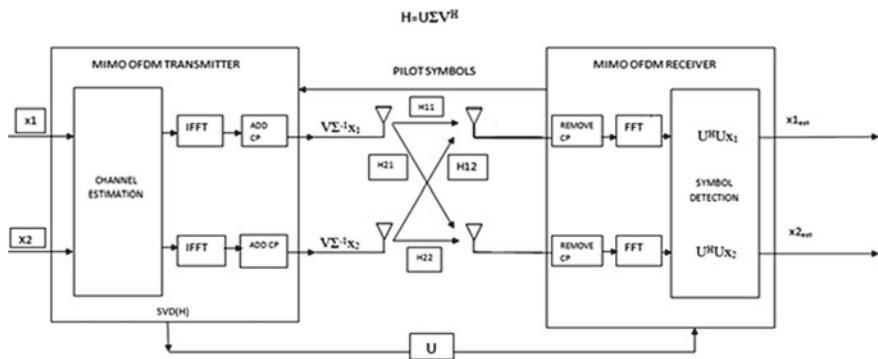


Fig. 4 Precoding using SVD

5.2 Alamouti-Based Precoding

The traditional Alamouti [1]-based precoding scheme is considered. One of the major problems with regard to the inefficiency of the Alamouti scheme is that it requires two-time slots to transfer two symbols. It requires transmission of symbols x_1 and x_2 in the first slot and $-x_2^*$ and x_1^* in the next. On combining the matrix values for symbols in two-time slots, the H matrix is obtained as

$$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \\ h_{12}^* & -h_{11}^* \\ h_{22}^* & -h_{21}^* \end{bmatrix}$$

The estimate of the symbols is

$$\begin{bmatrix} \widehat{x}_1 \\ x_2^* \end{bmatrix} = (H^H H)^{-1} H^H \begin{bmatrix} y_1^1 \\ y_1^2 \\ y_2^{2*} \\ y_2^{1*} \end{bmatrix}$$

The pseudoinverse can be performed by any of the estimation techniques.

5.3 QR-Based Precoding

A novel QR-based precoding is presented in [3]. However, this scheme can be applicable for systems with $N_t = 2 \times N_r$. An adaptation of this scheme provides a solution for a general case.

Once the transmit signal is computed for each antenna, a separate signal Q has to be transmitted to the receiver. On reception of the data a simple inverse operation of Q provides the reconstruction of the required symbol. This is described in Fig. 5.

The algorithm is as follows.

Algorithm – QR – Precoding

Inputs: $H \in C^{N_t \times N_r}$, data $x \in C^{N_r}$, unitary $U \in C^{N_t \times N_r}$, transmit power p

Outputs : Transmit signals $s \in C^{N_t}$

Start Compute QR decomposition $H=QR$

Partition $R = [R_1^H \ 0^H]^{(N_t-N_r) \times N_r}^H$,

Where $R_1 \in C^{N_r \times N_r}$

Compute $x' = R_1^{-H} x$ and $x^H = [x^H \ x'^H \ O^H]_{N_t-2N_r}$

Compute $P=QU$

Compute transmit signal $s = (1/N_t)^{1/2} P_x$

End

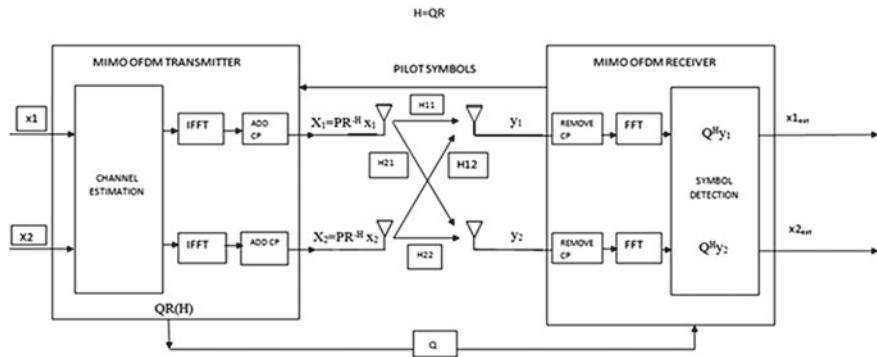


Fig. 5 Precoding using QR

6 Results and Discussion

The channel is assumed to be TDD and the simulation is done for 2×2 MIMO-OFDM system. The modulation schemes used are BPSK, QPSK, 16-QAM. The results are obtained for various channel estimation algorithms.

It is found that maximum likelihood estimation has the highest accuracy. The QPSK and 16-QAM modulation schemes use a lookup table at the receiver end to demap the signal into bits. As a result, in the practical transmission of data, appropriate power levels have to be considered. The voltages under consideration should not be very close because variations in noise level would lead to erroneous detection. However, large differences lead to the transmission of very high power which may cause heating of the mobile system. Table 1 provides the simulation parameters of the OFDM system.

Table 2 lists the MIMO-OFDM parameters for 16-QAM. From the graph, we infer that although OFDM provides a better BER performance when compared to the conventional FDM systems due to the orthogonality of subcarriers and introduction of cyclic prefix, a BER of 0.145 cannot be accepted for practical standards. This is the reason why we try to incorporate MIMO with the OFDM systems (Fig. 6 and Table 3).

Table 1 OFDM system parameters

Number of data bits	8×10^4
Bits/Symbol	1
Channel/Noise	Rayleigh/AWGN
BER for SVD precoding (at 13 dB)	3.8×10^{-4}
BER for Alamouti precoding (at 13 dB)	6×10^{-3}
BER for QR-based recoding (at 13 dB)	9×10^{-5}

Table 2 MIMO-OFDM parameters for 16-QAM modulation

Number of subcarriers	64
Channel	Rayleigh fading channel
OFDM bandwidth	20 MHz
Bandwidth/Carrier	3.125 MHz
OFDM symbol period	3.2 μ s
Cyclic prefix period	0.8 μ s
BER (at 15 dB)	0.145

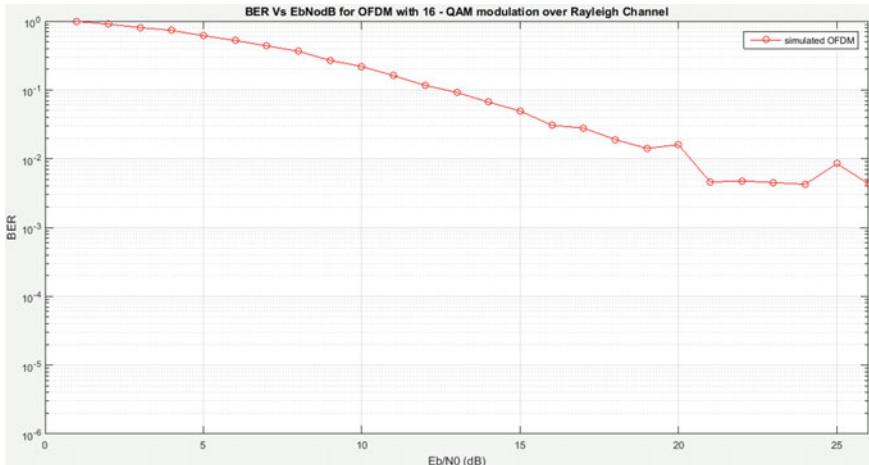


Fig. 6 BER versus SINR for OFDM modulation scheme

Table 3 MIMO-OFDM parameters for different precoding

Number of data bits	8×10^4
Bits/Symbol	1
Channel/Noise	Rayleigh/AWGN
BER for zero forcing (at 14 dB)	4×10^{-2}
BER for MMSE (at 14 dB)	2.8×10^{-2}
BER for Q less QR (at 14 dB)	2.2×10^{-2}
BER for MLD (at 14 dB)	1.2×10^{-4}

From Fig. 7 it is observed that the BER gradually increases with every modulation technique. This is because, for schemes like QAM, a single bit error leads to a burst error due to erroneous detection of the symbols. However, during practical scenarios, we use higher order modulation schemes due to the increase in data rate that they provide. Television systems employ transmission schemes like 256-QAM. This is also the reason why BPSK is used for transmission of pilot symbols. As pilots are used to ascertain the properties of the channel, high degree of accuracy is required which can be compromised in the data rate (Fig. 8).

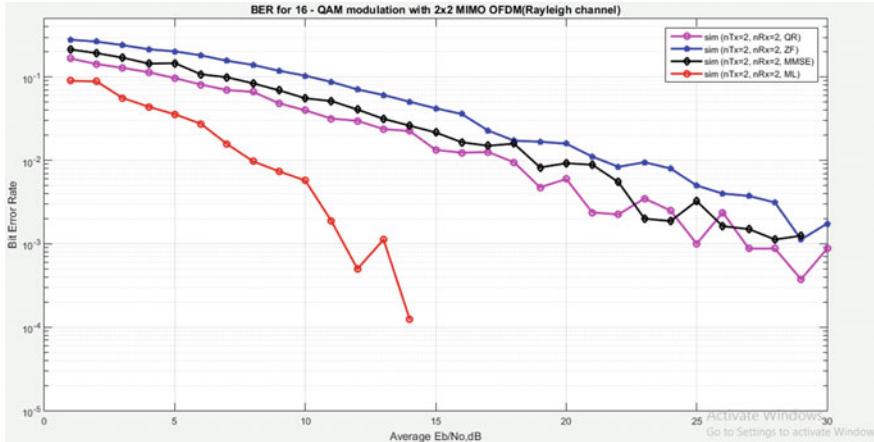


Fig. 7 BER versus SINR for MIMO-OFDM 16-QAM modulation

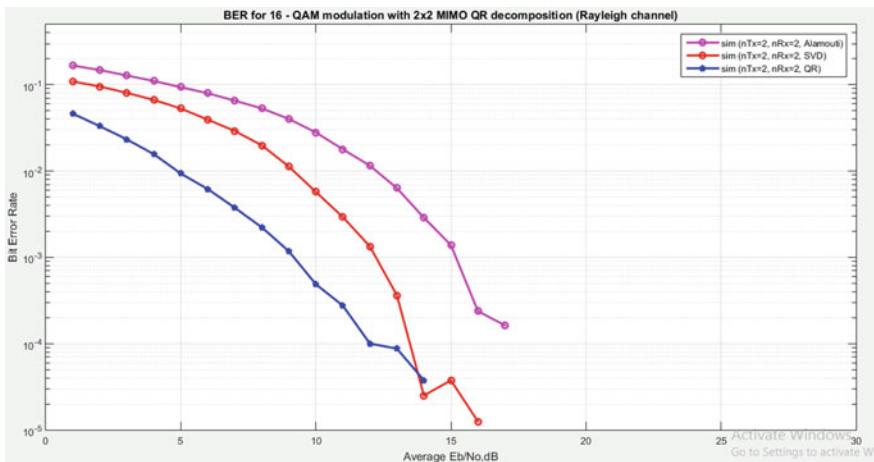


Fig. 8 BER versus SINR for different precoding techniques under 16-QAM

Having obtained the channel properties, precoding is employed to further decrease the BER. The proposed QR-based precoding offers a better performance against the SVD and conventional Alamouti Space–Time Coding Schemes. A theoretical BER of about 9×10^{-5} is achieved by this scheme (Figs. 9 and 10).

The capacity plot is obtained through the following formula for flat fading MIMO channel system.

$$C = \log(\det(I_{NT} + (\gamma/N_T)HH^H)) \quad (3)$$

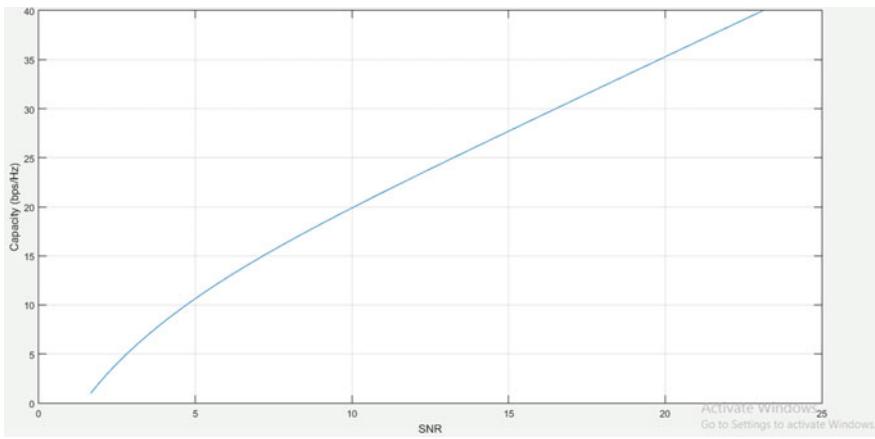


Fig. 9 Capacity versus SINR for 2×2 MIMO-OFDM system

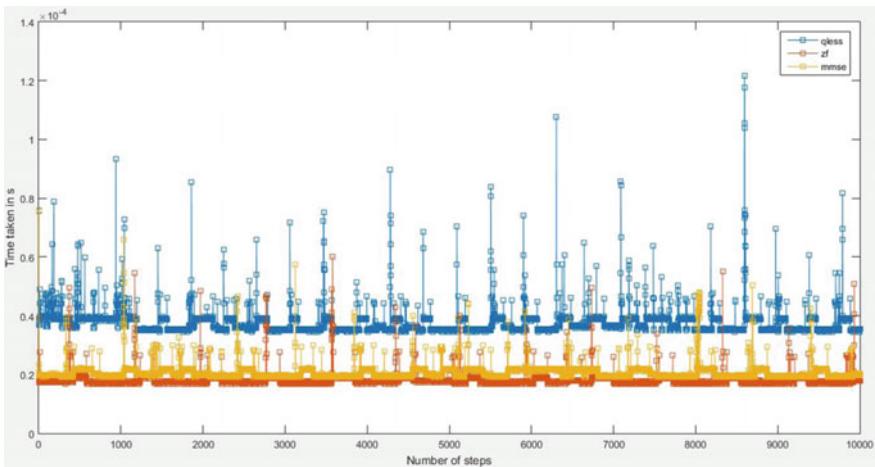


Fig. 10 Time complexity for the various iterations of Zero Forcing, MMSE and Q less QR-based decomposition

For, a practical SNR range of 20 dB, a capacity of 35 bps/Hz is obtained for our 2×2 MIMO-OFDM system.

7 Conclusion

The time complexity of the estimation techniques is obtained by finding the computational time of each technique. From the above graph, it is found that Zero Forcing algorithm takes the least time to compute, while ML detection takes the most. Also, MMSE takes lesser time to compute compared to Q less QR technique. However, considering the optimal performance, Q less QR technique offers better Bit Error Rate (BER). It can be observed that the use of precoding greatly reduces the BER but comes with an additional complexity due to an extra hardware required at the transmitter. From a BER of 2.2×10^{-2} by using the 16-QAM, it is possible to achieve a theoretical BER of 9×10^{-5} at a SINR of 14 dB. Of all the estimation techniques, Q less QR-based estimation offers an optimal BER. The Maximum Likelihood detection scheme suffers from the drawback that a large codebook is required at the transmitter, consuming more time than other schemes. The capacity of the MIMO-OFDM system is found out to be 35 bps/Hz at an SNR of 15 dB. The time complexity of Q less QR scheme is approximately equal to 0.4×10^{-4} s which is higher than that of other techniques.

References

1. Alamouti, S.M.: A simple transmit diversity technique for wireless communications. *IEEE J. Sel. Areas Commun.* **16**(8)
2. Gesbert, D.: From theory to practice: an overview of MIMO space-time coded wireless systems. *IEEE J. Sel. Areas Commun.* **21**(3) (2003)
3. Hochwald, B.M., Marzetta, T.L.: Unitary space-time modulation for multiple-antenna communications in rayleigh flat fading. *IEEE Trans. Inf. Theory* **46**(2), 543–564 (2000)
4. Jose, J., Ashikhmin, A., Whiting, P., Vishwanath, S.: Channel estimation and linear precoding in multiuser multiple-antenna TDD systems. *IEEE Trans. Veh. Technol.* **60**(5) (2011)
5. Lopes, A., Aalam, Z.: Performance analysis and hardware implementation of a 802.11n 4×4 MIMO OFDM transceiver using 16-QAM. *Int. J. Adv. Inf. Sci. Technol. (IJAIST)* **27**(27) (2014)
6. Molisch, A.: *Wireless Communications*. Wiley-IEEE Press (2005)
7. Rappaport, T.S.: *Wireless Communications: Principles and Practice*, 2nd edn. Pearson, India (2010)
8. Thiagarajan, G., et al.: Novel precoding methods for Rayleigh fading multiuser TDD-MIMO systems. In: *IEEE International Conference on Communications*, pp. 5586–5591, 10–14 June 2014
9. Thiagarajan, G., et al.: Novel transmit precoding methods for Rayleigh fading multiuser TDD-MIMO systems with CSIT and no CSIR. *IEEE Trans. Veh. Technol.* **64**(3), 973–984 (2015)
10. Tse, D., Viswanath, P.: *Fundamentals of Wireless Communication*. Cambridge University Press (2005)

Development of Sine Cosine Toolbox for LabVIEW



Shubham Maurya, Mohit Jain and Nikhil Pachauri

Abstract LabVIEW is one of the prominent programming platforms due to its simplicity and robustness towards hardware design. It finds application in different fields of engineering such as industrial automation, signal processing, instrumentation, control system design and so forth. LabVIEW has a rich ecosystem of toolboxes and libraries for simplifying the improvement in numerous zones of engineering, however, it lacks in optimization toolbox for advanced meta-heuristic algorithms. This is a serious deficiency as there is an astonishing improvement in the field of meta-heuristic algorithm in recent years. Thus, in this article, a potentially acclaimed meta-heuristic named as sine cosine algorithm (SCA) is designed in LabVIEW environment. The designed algorithm is benchmarked on a testbed of six non-trivial functions and compared to existing differential evolution (DE) optimizer in LabVIEW. Statistical analysis of results depicts superior accuracy and stability of SCA in comparison to DE. Further, SCA based PID optimizer is also implemented to show the applicability of the designed algorithm in solving real-world optimization problems.

Keywords Sine cosine algorithm · Optimization · LabVIEW

S. Maurya (✉)

Department of Electronics and Communication, GLA University, Mathura 281406, India
e-mail: er.shubham14@gmail.com

M. Jain

School of Automation, Banasthali Vidyapith, Vanasthali 304022, Rajasthan, India
e-mail: nsit.mohit@gmail.com

N. Pachauri

Department of Electrical and Instrumentation, Thapar Institute of Engineering and Technology,
Patiala 147004, India
e-mail: nikhil.pachauri@thapar.edu

1 Introduction

LabVIEW by National Instrument has been an important platform for the researchers from the past few years. It provides a graphical programming environment for real-time control and measurement purpose. It is extensively used in industries and academics due to its effortless dataflow programming and easy hardware interfacing. It facilitates the researchers with a number of inbuilt toolboxes, which are beneficial for several engineering applications. However, LabVIEW package consists of only one global optimization toolbox based on differential evolution (DE) algorithm. There is an inadequacy of proficient optimizers in LabVIEW. Initially, derivative-based approaches are used to solve engineering optimization problems. However, these methods are found inefficient to tackle complex non-linear problems with large dimension size. Recently, various nature-inspired meta-heuristic techniques are developed such as squirrel search algorithm [1], owl search algorithm [2] and improved crow search algorithm [3] to solve these complex problems efficiently. Another advanced meta-heuristic called as sine cosine algorithm (SCA) is proposed by Mirjalili [4], which is formulated on mathematical properties of sine and cosine functions. SCA is found efficient for solving sophisticated engineering problems like feature selection [5] and power system design [6]. This motivates the authors to design a novel optimization toolbox for LabVIEW based on SCA optimizer. The main contributions of this work are as follows:

1. A maiden attempt is made to design SCA based optimization toolbox in LabVIEW environment.
2. SCA based PID tuner is proposed to cope up with controller tuning problems.

The remaining paper is structured as follows: Sect. 2 explains the implementation of SCA in LabVIEW. Section 3 presents the result and discussion of the manuscript. Implementation of SCA based PID tuner is described in Sect. 4. Conclusion of the work is discussed in Sect. 5.

2 Design of SCA Optimizer in LabVIEW

SCA is inspired from the fundamental characteristics of sine and cosine functions. It utilizes the range of sine and cosine functions to maintain a balance between exploration and exploitation [1]. Development of the SCA optimizer includes the implementation of different Virtual Instruments (VIs), which handle the following tasks:

- (a) Random initialization of solutions.
- (b) Evaluation of each solution and store the best one out of them.
- (c) Update the range of sine and cosine functions.
- (d) Update the target point.
- (e) Repeat the process from step (b) until the stopping criteria is fulfilled.

The purpose of each VI is discussed as follows:

2.1 Generate Initial Population.VI

This VI is used to generate random initial solutions on the basis of user-defined parameters like upper bound (UB), lower bound (LB), dimensions (D) and no of search agents (N). The following equation is used to implement this VI (Fig. 1).

$$X = \text{rand}(N, D) \times (UB - LB) + LB \quad (1)$$

2.2 Optimal Solution.VI

The function of this VI is to examine the fitness of each solution by employing user-defined objective function. It has two dimensional randomly generated solution array as an input and returns one-dimensional array of fitness corresponding to each solution (Fig. 2).

2.3 Position Update.VI

This VI (Fig. 3) is a pillar for SCA toolbox and employed to update the position of current solution set. The following equations are used at any t th iteration to update the i th dimension of present solution (X'_i):

$$X_i^{t+1} = \begin{cases} X_i^t + r_1 \times \sin(r_2) \times |r_3 P_i^t - X_i^t|, & r_4 < 0.5 \\ X_i^t + r_1 \times \cos(r_2) \times |r_3 P_i^t - X_i^t|, & r_4 \geq 0.5 \end{cases} \quad (2)$$

$$r_1 = a - t \frac{a}{T} \quad (3)$$

Fig. 1 Generate initial population.VI

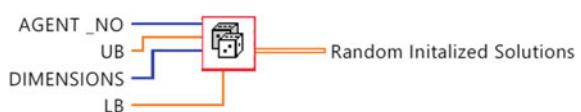


Fig. 2 Optimal solution.VI

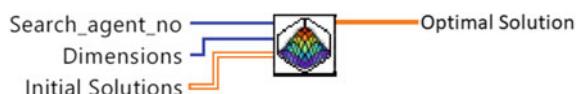
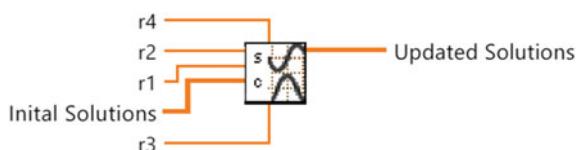


Fig. 3 Position update.VI



where a is a constant, T is total no of iterations and P_i is i th component of target solution. The parameter r_1 controls the direction of next location which lies in the region between the present and target solution or exterior to it. r_2 decides the distance of movement near or far away from the target. The parameter r_3 assigns random weight to target in order to randomly emphasize ($r_3 > 1$) or de-emphasize ($r_3 < 1$) the effect of desalination in describing the distance. r_4 provides probabilistic switching among sine and cosine factors (Eq. 2).

3 Result and Discussion

The proposed SCA based optimization toolbox is developed in LabVIEW™ version 14.0 on PC having 64-bit operating system, Intel® Core™ i3-4005U 1.70 GHz processor and 4 GB RAM. Figures 4 and 5 display the front panel and block diagram for designed SCA Optimizer. The performance of proposed toolkit is validated on unimodal and multimodal benchmark functions. DE algorithm is also employed on these benchmark functions for comparative study. Both the algorithms are executed for 10 independent runs with maximum 1000 iterations and 20 search agents. Statistical performance measures like mean and standard deviation are recorded and presented in Table 1. The values below 10E-10 are treated as zero, i.e. optimal value. It is revealed that SCA attains optimal values for each test functions while DE traps in local minima for f_1, f_4 and f_6 functions. However, DE reaches to near optimal region for f_2, f_3 and f_5 functions. Moreover, SCA offers less standard deviation in comparison to DE. Thus, SCA proves to be more accurate, stable and consistent

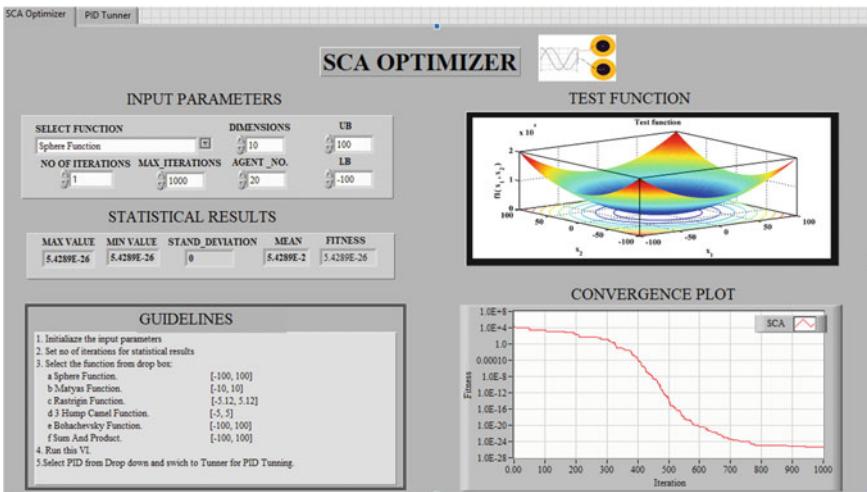


Fig. 4 The front panel of SCA optimizer

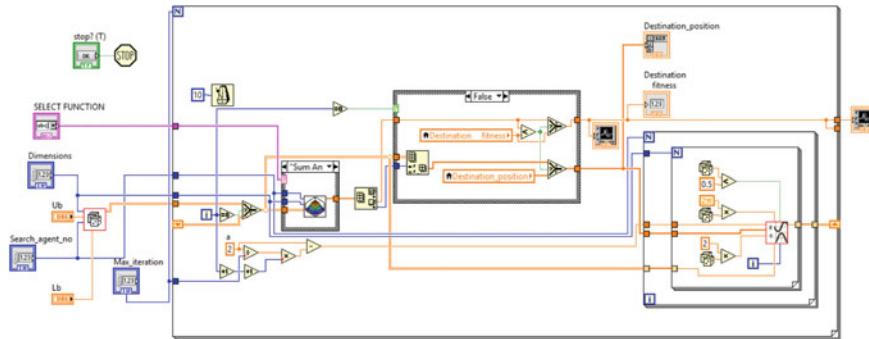


Fig. 5 The block diagram of SCA optimizer

Table 1 Comparison of mean and standard deviation of SCA and DE for benchmark functions

Function	Mean		Standard deviation	
	SCA	DE	SCA	DE
Sphere (f_1)	3.75628E-26	0.019714211	6.70149E-24	0.027442346
Sum and product (f_2)	7.22847E-29	2.918046E-4	5.3025E-24	2.734758E-4
Matyas (f_3)	3.32732E-34	3.84362E-5	1.02854E-29	3.62578E-5
Rastrigin (f_4)	0	0.014388465	2.7862E-13	0.029719132
3-hump camel (f_5)	6.46429E-60	9.37848E-5	2.24765E-52	0.00010913
Bohachevsky (f_6)	0	0.745367	0	0.193716

as compared to the existing DE optimizer in LabVIEW. Further, the practical utility of the proposed SCA optimization toolbox is tested by employing it for parametric tuning of PID controller.

4 Generalized SCA Based PID Tuner

Proportional, integral and derivative (PID) controllers are widely utilized by the researchers because of its robustness and reliability. They provide effective solutions to various complex non-linear processes such as heat exchanger, chemical reactors, etc. The basic equation of PID controller in Laplace domain is given by [7, 8]

$$M(s) = E(s)\{K_p + K_i/s + sK_d\} \quad (4)$$

where $M(s)$ is the manipulated variable and $E(s)$ is error between set-point and output of the plant. K_p , K_i and K_d are proportional, integral and derivative gains, respectively, and their appropriate values lead to optimal performance of the plant. Classical PID tuning methods are generally based on fixed performance measure and

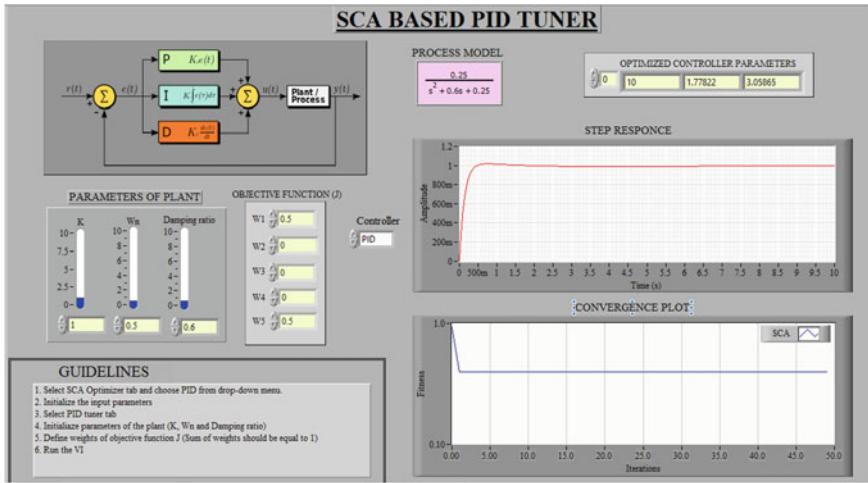


Fig. 6 Front panel of generalized PID tuner

there is no flexibility to consider multiple performance measures simultaneously. Thus, in the present work, a flexible PID tuning scheme is developed in which an objective function is designed by considering a weighted combination of various performance indices like integral absolute error (E_{IAE}), integral square error (E_{ISE}), integral time-weighted absolute error (E_{ITAE}), settling time t_{sett} and percentage of overshoot (OS) (Eq. 5)

$$J = w_1 \times E_{IAE} + w_2 \times E_{ISE} + w_3 \times E_{ITAE} + w_4 \times t_{sett} + w_5 \times OS. \quad (5)$$

where the values of w_1 , w_2 , w_3 , w_4 and w_5 are selected by control system designer such that (Fig. 6)

$$w_1 + w_2 + w_3 + w_4 + w_5 = 1 \quad (6)$$

5 Conclusions

The present research work is focused on the design and development of LabVIEW toolbox for a recently claimed sine cosine algorithm. A comparative study is performed between SCA and DE on the testbed of six benchmark functions. Results depict the superior performance of SCA as compared to the existing DE optimizer on all benchmark functions. Hence, it is concluded that the developed SCA optimizer is reliable, consistent, stable and provides optimal solution in almost every run for a

given optimization problem. Further, SCA is also found effective in PID controller tuning. In future, the present work may be extended for constrained optimization problems with improved versions of SCA claimed in the literature.

References

1. Jain, M., Singh, V., Rani, A.: A novel nature-inspired algorithm for optimization: squirrel search algorithm. *Swarm Evol. Comput.* (2018). <https://doi.org/10.1016/j.swevo.2018.02.013>
2. Jain, M., Maurya, S., Rani, A., Singh, V.: Owl search algorithm: a novel nature-inspired heuristic paradigm for global optimization. *J. Intell. Fuzzy Syst.* **34**(3), 1573–1582 (2018)
3. Jain, M., Rani, A., Singh, V.: An improved crow search algorithm for high-dimensional problems. *J. Intell. Fuzzy Syst.* **33**(6), 3597–3614 (2017)
4. Mirjalili, S.: SCA: a sine cosine algorithm for solving optimization problems. *Knowl. Based Syst.* **96**, 120–133 (2016)
5. Hafez, A.I., Zawbaa, H.M., Emary, E., Hassanien, A.E.: Sine cosine optimization algorithm for feature selection. In: 2016 International Symposium on INnovations in Intelligent SysTems and Applications (INISTA), pp. 1–5. IEEE (2016)
6. Attia, A.F., El Sehiemy, R.A., Hasanien, H.M.: Optimal power flow solution in power systems using a novel Sine-Cosine algorithm. *Int. J. Electr. Power Energy Syst.* **99**, 331–343 (2018)
7. Pachauri, N., Singh, V., Rani, A.: Two degree of freedom PID based inferential control of continuous bioreactor for ethanol production. *ISA Trans.* **68**, 235–250 (2017)
8. Pachauri, N., Rani, A., Singh, V.: Bioreactor temperature control using modified fractional order IMC-PID for ethanol production. *Chem. Eng. Res. Des.* **122**, 97–112 (2017)

A Survey on Big Data in Healthcare Applications



M. Ambigavathi and D. Sridharan

Abstract The large volume of healthcare data continues to mount every second, making it harder and very difficult to find any form of useful information. Recently, big data is changing the traditional way of the data delivery system into valuable insights, especially in the healthcare industry. It provides a lot of benefits in the healthcare sector to detect critical diseases at the initial stage and deliver better healthcare services to the right patient at the right time. It has provided tools to accumulate, manage, analyze, and assimilate large volumes of disparate, structured, and unstructured vital data rapidly produced by the various healthcare information storage systems. However, there are several issues to be addressed in the current health data analytics platforms that offer technical methods such as vital data collection, aggregation, process, analysis, visualization, and sharing. Due to lack of detailed analysis in the existing research works, this paper examines the most significant characteristics of big data analytics in health care, various data sources and its data types, five R's of big data analytics, and then briefly discusses the recent open research challenges with future directions.

Keywords Big data · Big data analytics · Healthcare data sources · Healthcare data types · Healthcare applications

1 Introduction

In the past decades, the healthcare sectors had been used the conservative method for diagnosis and medical treatment of the patients. Since the majority of medical practitioners or physicians are dependent on their individual levels of information

M. Ambigavathi (✉) · D. Sridharan
Department of Electronics and Communication Engineering, CEG Campus,
Anna University, Chennai, India
e-mail: ambigaindhu8@gmail.com

D. Sridharan
e-mail: sridhar@annauniv.edu

and skills to provide suggestions and feedback about the symptoms and effects of illness. This may result in incorrect, imprecise, and non-patient-centered care [1, 2]. Big data is an emerging technology to modernize the conventional healthcare system and move the current healthcare industry forward on several fronts [3]. The data from different sources such as mobile phones, body area sensors, patients, hospitals, researchers, healthcare providers, and organizations are currently generating the immense volume of medical data. The amassed data from various sources will be in dissimilar forms, including electronic health records (EHRs), medical imaging (MI), genomic sequencing (GS), clinical records (CR), pharmaceutical research (PR), wearable, and medical devices (MD). These big health data are specially stored in a medical server (MS), clinical database (CDB), and other clinical data repositories (CDR) for next level analysis [4]. The storage infrastructures are primarily [5] used to store, process, analyze, manage, and retrieve the huge amounts of data in order to make it easier for the people. Therefore, it is not only providing the information to understand the symptoms, illness, and therapies but also to alert, predict outcomes at initial stages and make the right decisions.

Big data analytics tool [6] is a new paradigm, mainly designed to analyze, manage and precisely extract the useful information from large volumes of data sets that are very similar to a particular patient in a very short time. Some of the common software tools are used as part of this advanced analytics strategy such as predictive analytics, data mining, text analytics, and statistical analysis. Moreover, this modern technological-based analytics method transforms the healthcare industry entirely to take the right decision for the right patient at the right time [7]. These factors are motivated us to examine the characteristics, various data sources and data types, five R's of big data analytics in healthcare applications. The main contributions of this article are outlined as follows:

- To study the most essential five V's of big data characteristics in health care.
- To present the diverse data sources of big health data analytics.
- To investigate the benefits of big data analytics in health care.
- To discuss the open research challenges with feasible solutions.

The remainder of this article is outlined as follows: Sect. 2 provides the characteristics of big data analytics in health care. Section 3 discusses the various medical data sources and its data types. The five R's of big data analytics is presented in Sect. 4. Open research challenges of big data analytics in health care is reported in Sect. 5. Lastly, this paper concludes with future directions in Sect. 6.

2 Characteristics of Big Data Analytics

This section describes the evolution of V's and characteristics of big data in the healthcare applications [8]. The characteristics of big health data are mainly associated with various issues, namely capture, cleaning, curation, integration, storage, processing, indexing, search, sharing, transfer, mining, analysis, and visualization.

Several researchers at different periods are recommended a wide range of V's for different applications. Also, some researchers also suggest that it will be increased up to 100 V's in the near future for an efficient big data analytics. The most important characteristics of bid data in the healthcare application are summarized as follows.

2.1 Volume

Volume is the amount of data continuously generated from various health monitoring sources. Moreover, the size healthcare-related data may vary from octet to Geopbytes, the rate including Kilobytes (1024 bytes), Megabyte (1024 kB), Gigabyte (1024 MB), Terabytes (1024 GB), Petabytes (1024 TB), Exabytes (1024 PB), Zettabytes (1024 EB), Yottabytes (1024 ZB), Brontabytes (1024 YB), and Geopbytes (1024 BB). Further, the estimated volume of healthcare data will be raised to 2314 EB by 2020 based on the Stanford medicine 2017 health trends report.

2.2 Variety

The data generated from various sources are stored in different types and formats such as structured, semi-structured, unstructured data and quasi structured. This variety of data is closely analyzed by using analytical tools, and delivered the right solution for a specific patient [9].

2.3 Velocity

The high rate of medical data is created, captured, refreshed at a high speed and shared in a very short time from milliseconds to hours. These data are mostly delay-sensitive and often in a correct format, but not always even if it is updated on a regular basis. Thus, the big health data must be analyzed, compared with recently generated data and retrieved to give more accurate decisions.

2.4 Veracity

This is a combination of both data consistency and data trustworthiness. The collected big health data should be accurate, completeness, unambiguous, righteous and confidential. In fact, the analysis of data accuracy fully depends on the veracity of the source data. Therefore, high-quality data is necessary to provide effective results using data analytics.

2.5 Veracity

It is similar to veracity but validity specifies the data quality in terms of accurate and correctness from the distributed and unclean medical data collections. So, the thorough analysis of healthcare data is essential for right and valid predictions.

3 Big Health Data Sources

The massive growth of data in the healthcare industry from diverse data sources has been categorized into two types: internal and external sources [10] and [11]. The data is also gathered in different formats (i.e., text, flat files, .csv, relational tables, images, audios, and videos, respectively) and placed at several locations (i.e., healthcare provider's sites as well as the physician's sites) and other applications like medical information transaction processing applications, databases, etc. [12]. Moreover, these data sources are mainly used by the medical professionals to help him or her to make correct decisions and provide appropriate treatment in an efficient way.

3.1 External Sources and Data Types

Web and social media data. This type of data in the healthcare sector is generated from the Facebook sharing, Twitter, LinkedIn, healthcare blogs, health plan websites, news feeds, healthcare professional's personal websites, etc.

Healthcare administrative data. The administrative data such as billing, patient scheduling for diagnosis, and other insurance claims will come under this category.

Device-to-device data. The medical monitoring devices may create the heterogeneous amount of vital parameters from sensors, remote devices, and other storage devices.

Health-related publication and clinical reference data. These data may consist of text-based publications such as health-related journals or articles, clinical research materials, clinical text-based guidelines and health product specific information on drug, etc.

Biometric data. This group of data may be classified as fingerprints, handwriting, iris scans, etc., which are used to smart identification of the patient in order to secure access to sensitive health records and to support with patient registration process.

Other important data: This would also include the feedback, content from the portal or individual health records (IHR) between the patient and the healthcare providers.

3.2 Internal Sources and Data Types

Human-generated data. The multimodal sensed data are collected from wearable or implantable monitoring sensor devices, ambulatory devices, and smartphones from indoor (i.e., home) or in outdoor (i.e., hospitals) that requires to aggregate, process, analyze the type of vital parameters including blood pressure, blood glucose, respiration rate, heart rate, etc.

Clinical transaction data. The clinical professionals also share medical information to remote monitoring end in different formats to evaluate the report and discuss with other entities for further diagnosis. This category of data may also contain X-rays, other medical images (i.e., ultrasound or computed tomography scan), medical records (i.e., from hospitals, medical research centers, government agencies, and insurance companies), physician's notes, emails, and paper documents.

Genomic data: The genomic data may also include the genotyping, gene expression and DNA sequences, etc. One of the major challenging tasks is to collect, analyze, and classify the data on these genes and observe the genetic level to the physiological level of a human being.

4 Five R's of Big Health Data Analytics

The recent advent of disruptive technology and the pace of big data analytics in the healthcare industry are incredibly fast with the astronomical increase in the degree of electronic healthcare records [13]. This new innovative big data analytics technique has a valuable impact and views on healthcare applications through five different R's as follows [14].

4.1 Right Living

Right living helps the patient to live healthier and easier by offering early detection and diagnosis through real-time monitoring devices. If they receive any abnormal status in their physiological parameters that create alerts to the patients, physicians, and family members in order to take necessary action based on the observed information. Also, elderly people can now take care of their health by analyzing the data mining results to make better decisions to improve their health conditions. The right living progress helps the patients to get necessary suggestions or preventive measures from the doctors in an effective way by minimizing readmission procedures thereby reducing individuals cost.

4.2 Right Care

The big data analytics ensures that the patients to take the right decisions on their treatment at the right time. It is also required that all medical professionals should have the same medical information and the same goals while deciding any treatment based on the historical data (i.e., set of similar patients, symptoms, and drugs, respectively) to avoid duplicated and incorrect information. Moreover, this big data solution provides the evidence-based customized line of therapy for a patient based on the location, climate, routine life, order of the past occurrences, hypersensitivity, genomic data, and so on.

4.3 Right Provider

Based on the outcome of the analyzed data, the healthcare providers must help to identify, predict and prevent the diseases of the right patients. The big data analytics results allow healthcare providers to support effective treatment. In addition, the hospitals (i.e., local, state or government) are set up according to the norms and standards established by the Indian Medical Council (IMC) and Indian Public Health Standards (IPHS). Furthermore, healthcare providers should introduce new health insurance policies and plans to reduce healthcare cost. This standard is basically comprised of five fundamental aspects such as functional program, space, manpower, equipment, and building requirements. These standards should be periodically inspected in order to help the government in taking necessary actions and to provide effective medicine to the right patients.

4.4 Right Innovation

This approach is mainly used to find out new diseases, new treatment, and new tools for medical data analysis. Sometimes, the health information exchange (HIE) can increase the burden of the fraudulent activities. Therefore, the innovative predictive models can be used to predict and prevent fraud at the point of transactions. In evidence-based medicine, the caregivers might not provide sufficient evidence based on the smaller medical data sets. In such a case, they must require to verify if there is variation within the best available clinical data sets using different novel analytical mechanisms. The healthcare industry should need to focus more on the early detection of security threats and fraud in health care by improving innovative skills and engines in health care.

4.5 *Right Value*

The medical professionals and patients must understand the value of the future healthcare system with the purpose of constantly enhancing the value and its quality by using the big data analytics. Moreover, these analytical based solutions will reduce the total healthcare expenditure as well as the large amount of expenditure allocated by governments or other healthcare organizations for care-related activities to the patients such as medical errors, hospital admissions and readmissions, hospital efficiency, critical care services, etc. Recently, the increasing number of smart sensors, smartphones, and other smart personal assistant devices are also reducing the healthcare cost in the patient end. The recent startups for smart health care are inspected from various sources and acquired a lot of information for healthcare organizations to modernize its conventional system into a smart healthcare system.

5 Open Research Challenges

This section describes the most significant challenges of big data analytics in the medical industry. The below highlights different crucial factors that all healthcare industry should take into consideration before the medical practitioners make a decision or form an opinion to the right patient. Most of the data cleaning processes are still facing a lot of challenges and offer automated scrubbing analytical tools [15] that use logic rules to compare, contrast, and correct large medical datasets. Therefore, these tools will be necessary to adapt the current technology in order to make the actual data sets faultless. In addition, the data quality [16] fully depends on the data accuracy so that the big data analysis techniques can also be used to perform the data cleansing process. However, the missing data, inconsistent data, data conflicts, and duplicates data will result in data quality challenges [17]. These incompleteness and errors must be eliminated during the data analysis. It is necessary for healthcare organizations to gain important insights from the big health data analysis. Further, the volume of more healthcare data eventually leads to a risk of exposure of the data, making it highly vulnerable [18–20]. Accordingly, it is very essential for analysts and data scientists to consider these security issues and deal with the data in such a way that will not lead to the disruption of privacy.

6 Conclusion

This article presents the recent investigation on big data analytics in the healthcare industry. This paper explicates the potential characteristics, various data sources and data types, and then different types of R's for right healthcare services. Some of the open research challenges and feasible solutions are highlighted in order to

promote the patient's health, reduce healthcare cost, enhance treatment, and improve healthcare value and quality. With the help of analytics tools, data scientists can able to aggregate medical information from both internal and external sources. Ultimately, physicians can be alerted to do their treatment and reach out to patients in an efficient way. This massive volume of huge healthcare data eventually leads to a risk of exposure of the data, makes it highly vulnerable. Thus, it is very essential for analysts and data scientists to consider these security issues and deal with the data in such a way that will not lead to the disruption of privacy. In order to predict and take more accurate decisions, the machine learning algorithms are essential in the modern healthcare sectors, thus these algorithms for big data analytics will be discussed in future.

References

1. Kaur, P., Sharma, M., Mittal, M.: Big data and machine learning based secure healthcare framework. In: International Conference on Computational Intelligence and Data Science (ICCIDIS 2018), Elsevier Procedia Computer Science, vol. 132, pp. 1049–1059 (2018)
2. Singh, A.N.: Customized biomedical informatics big data analytics for personalized genomics. Springer J. Big Data Anal. 1–12 (2018)
3. Khan, N., Yaqoob, I., Hashem, I.A.T., Inayat, Z., Ali, W.K.M., Alam, M., Shiraz, M., Gani, A.: Big data: survey, technologies, opportunities, and challenges. Sci. World J. 1–18 (2014)
4. Wang, Y., Kung, L.A., Wang, W.Y.C., Cegielskid, C.G.: An integrated big data analytics-enabled transformation model: application to health care. Elsevier J. Inf. Manag. **55**, 64–79 (2018)
5. Godfrey, A., Hetherington, V., Shum, H., Bonato, P., Lovell, N.H., Stuart, S.: From A to Z: wearable technology explained. Elsevier J. Matur. 113, 40–47 (2018)
6. Raghupathi, W., Raghupathi, V.: Big data analytics in healthcare: promise and potential. Health Inf. Sci. Syst. 1–10 (2014)
7. Pramanik, M.I., Lau, R.Y.K., Demirkhan, H., Azad, M.A.K.: Smart health: big data enabled health paradigm within smart cities. J. Expert Syst. Appl. **87**, 370–383 (2017)
8. Abouelmehdi, K., Beni-Hessane, A., Khaloufi, H.: Big healthcare data: preserving security and privacy. J. Big Data 1–18 (2018)
9. Das, N., Das, L., Rautaray, S.S., Pandey, M.: Big data analytics for medical applications. Int. J. Mod. Educ. Comput. Sci. **2**, 35–42 (2018)
10. Herland, M., Khoshgoftaar, T.M., Bauder, R.A.: Big data fraud detection using multiple medicare data sources. J. Big Data 1–12 (2018)
11. El Aboudi, N., Benhlima, L.: Big data management for healthcare systems: architecture, requirements, and implementation. Adv. Bioinform. 1–10 (2018)
12. Rahman, F., Slepian, M.J.: Application of big-data in healthcare analytics prospects and challenges. In: IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), pp. 1–4 (2016)
13. Acharjya, D.P., Kauser Ahmed, P.: A survey on big data analytics: challenges, open research issues and tools. Int. J. Adv. Comput. Sci. Appl. **7**(2), 511–518 (2016)
14. Asri, H., Mousannif, H., Al Moatassime, H., Noel, T.: Big data in healthcare: challenges and opportunities. In: International Conference on Cloud Technologies and Applications (CloudTech), pp. 1–7 (2015)
15. Alkhatib, M.A., Talaei-Khoei, A., Ghapanchi, A.H.: Analysis of research in healthcare data analytics. In: Australasian Conference on Information Systems, pp. 1–16 (2015)

16. Reddy, S.S.R.D., Ramanadham, U.K.: Big data analytics for healthcare organization, BDA process, benefits and challenges of BDA: a review. *J. Adv. Sci. Technol. Eng. Syst.* **2**(4), 189–196 (2017)
17. Minor, L.: Harnessing the power of data in health. *Stanford Med.* 1–18 (2017)
18. Zhang, L., Wang, H., Li, Q., Zhao, M.-H., Zhan, Q.-M.: Big data and medical research in China. *Br. Med. J.* 1–3 (2018)
19. Kliment Ohridski, St.: Big data analytics in medicine and healthcare. *J. Integr. Bioinform.* 1–5 (2018)
20. Ram Mohan Rao, P., Murali Krishna, S., Siva Kumar, A.P.: Privacy preservation techniques in big data analytics: a survey. *J. Big Data* 1–12 (2018)

Solar-Assisted Cloud-Enabled Intelligent Mini Refrigerator



Nishi Khetwal, Basant Singh Bhaskar, Udayveer Mittal, Sourav Thakur, Vindhya Devalla, Adesh Kumar and Vivek Kaundal

Abstract In this paper, a solar-assisted cloud-enabled intelligent mini refrigerator has been proposed for the utilization of the wasteover food and to avoid food wastage. With the help of this uniquely proposed solution, the food which is leftover or not being consumed in the restaurants, societies, flats, hostels, and colleges can now be provided to the needy people. This solution will provide the availability of the food to the cloud-enabled app as soon as the food is detected in the smart refrigerator. So, the NGOs and community that work for the poverty-stricken people around a specific part of the country can keep track of the waste food in the various parts of the locality through the developed app and can judiciously distribute the same.

N. Khetwal · S. Thakur · V. Kaundal (✉)

Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies, Dehradun, India

e-mail: vkaundal@ddn.upes.ac.in

N. Khetwal

e-mail: nishi.0710@gmail.com

S. Thakur

e-mail: souravnariyal31@gmail.com

B. S. Bhaskar · U. Mittal

Department of Mechanical Engineering, University of Petroleum and Energy Studies, Dehradun, India

e-mail: basantsbhaskar@gmail.com

U. Mittal

e-mail: uvm99@outlook.com

V. Devalla

Department of Aerospace Engineering, University of Petroleum and Energy Studies, Dehradun, India

e-mail: vindsdevalla@yahoo.com

A. Kumar

Department of Electrical and Electronics Engineering, School of Engineering, University of Petroleum and Energy Studies, Dehradun, India

e-mail: adeshkumar@ddn.upes.ac.in

Keywords Solar photovoltaic · IoT · Cloud · Food wastage

1 Introduction

“Throwing away food is like stealing from the table of those who are poor and hungry” [1]. Hunger issue is the biggest issue that we are facing in the present time [2]. Children are the most visible victims of undernutrition. A principal problem is that many people in the world still do not have sufficient income to purchase (or land to grow) enough food or access to nutritious food. Food wastage cripples a country’s economy to an extent that most of us are unaware.

The proposed solution of intelligent mini refrigerator will be well accepted to the visible victims of food. Intelligent mini refrigerator will have a number of sensors that will provide the real-time availability of food in the fridge as well as about the type of food and its quantity. All these details will be accessible through the cloud-enabled app. The proposed system is also safe and secure with the smart lock provided that will give access to the specific NGOs and community by only through entering the unique code given in the app. The device can be kept at the entry gate of the desired places so that seekers can get food easily. The system will upload the information of food to the internet through the Wi-Fi. It will be powered by a solar panel as well as mains electricity. In the current scope of the paper, solar-driven intelligent mini refrigerator with cloud enabled has been discussed.

2 Objectives of Solar-Assisted Cloud-Enabled Intelligent Mini Refrigerator (SIMR)

With SIMR, the following two major objectives can be met:

- a. Hunger issue.
- b. Food wastage.

Interlinking the above issues and solving them simultaneously using SIMR by:

- a. Developing an IOT-based system.
- b. Designing a cloud-based app that will work on software as a service platform (SaaS).

3 Importance of the SIMR Prototype

To understand the magnitude of the issue, let us see a few facts from the Indian population [3]:

- a. 14.9% of our population is undernourished,
- b. 195.9 million people go hungry every day,
- c. 21.0% of children under 5 are underweight,
- d. 38.4% of children under 5 years of age are stunted,
- e. 1 in 4 children malnourished.

The proposed solution aims to bring a cost-effective way out to some of the major problems related to hunger and food wastage. It is saving our resources, which are wasted in making and producing food. It will save the life of people who die because of hunger.

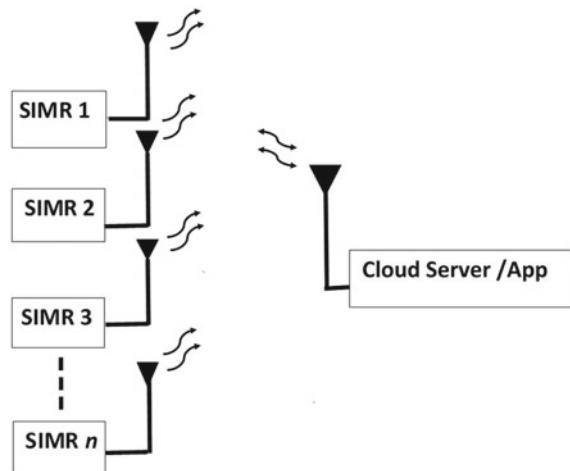
SIMR needs to be installed in common places like society flats, hotels, and hostels where food is wasted in huge amount. With the help of local NGOs and with the collaboration of hostels, restaurants, and residential societies food wastage can be checked and diverting the same to needy ones on a regular basis.

4 Methodology

4.1 Block Diagram

Figure 1 shows the generalized block diagram of SIMR. Here n number of SIMR devices communicated to server with its own unique API key, which remove the data interference among the multiple SIMR.

Fig. 1 Generalized system diagram SIMR



4.2 System Components (Shown in Fig. 2)

- a. Solar photovoltaic panel [4]: SIMR prototype is solar powered which will not only save the electricity but will also make it energy efficient and eco-friendly (details in Table 1).
- b. IoT [5]: System will be Wi-Fi enabled that will easily connect with the Wi-Fi network and fetch as well as process the information of the food to the dedicated app. The TCP/IP microcontroller ESP8266 will be used to enable Wi-Fi protocol (details in Table 1).
- c. Load Cell [6]: It is installed with different keys will tell different states of refrigerators whether full, empty or half-filled with corresponding weight (details in Table 1).
- d. Keypad and display: In residential societies or hotels and their places, the smart refrigerator will open to only to the registered customers, i.e., NGOs. In which, the specific lock code will be accessible to the app. To assist, the same keypad and display are attached.
- e. GPS: Geolocation of the various SIMR systems around different location will be tracked through GPS (details in Table 1).

The developed prototype is shown in Fig. 3.

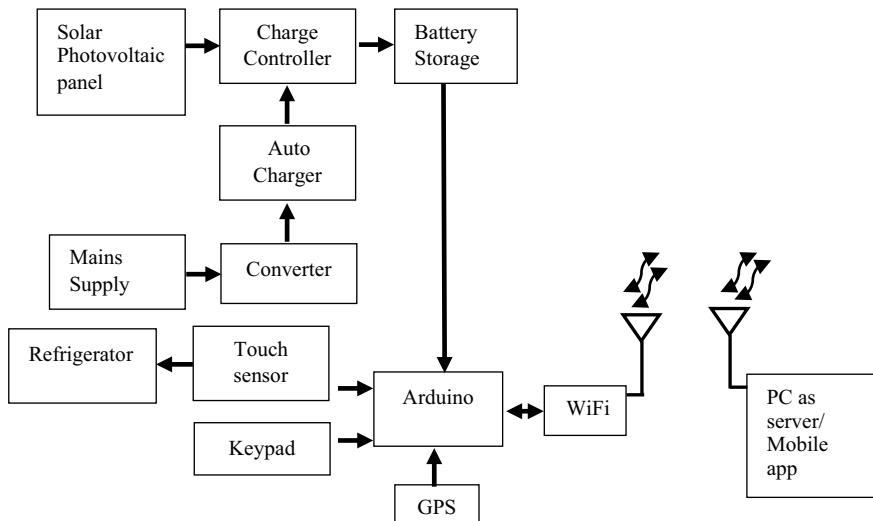


Fig. 2 SIMR 1 smart device block diagram

Table 1 Specifications of system components

Solar photovoltaic panel	IoT	Load cell	GPS
Size: 4 cm × 3.3 cm × 1 cm	Protocol: 802.11 b/g/n	Model: YZC-131	U-blox/u-blox NEO-6 M GPS module with antenna and build-in EEPROM
Output: 5 V	Wi-Fi direct (P2P), soft-AP	Rated load: 5 kg	Compatible with APM2 and APM2.5
Input: 6–20 V	Integrated TCP/IP protocol stack	Rated output: 1.00.15 mV/V	Interface: RS232 TTL
Output current: 2 A	Integrated TR switch, LNA, power amplifier and matching network	Nonlinear: 0.03%F.S	Power: 3–5 V
Operating temperature: −40 to +85°C	Integrated PLL, regulators, and power management units	Hysteresis: 0.03%F.S	Baudrate default: 9600 bps
	+19.5 dBm output power in 802.11 b mode	Repeatability: 0.03%F.S	
	Integrated temperature sensor	Creep (5 min): 0.05%F.S	
	Supports antenna diversity		
	Power down leakage current of <10 uA		

4.3 App Development

The complete information from the refrigerator will be stored in the cloud and will further transferred to the local NGOs which will collect the food. Mobile app will make the work of local. In the proposed app, as shown in Fig. 4, meter depicts the fridge status (empty, half-filled or full).

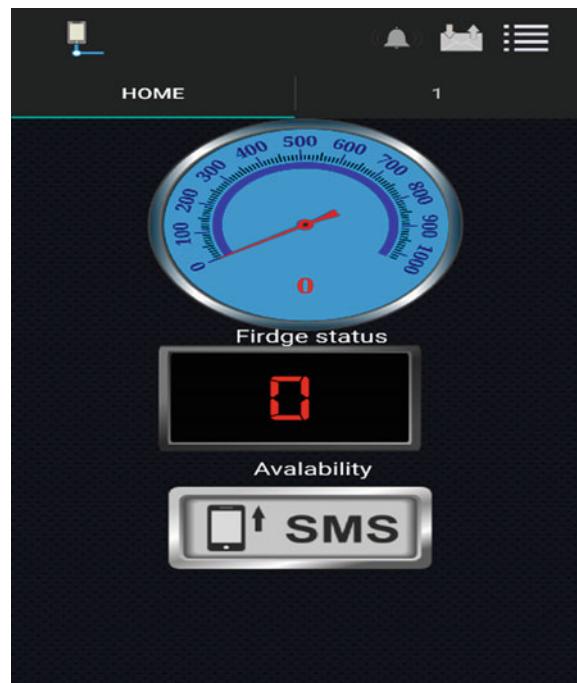
5 SIMR Energy Requirement

SIMR is solar-assisted refrigerator, which is able to run on solar as well as on mains supply. SIMR is also equipped with an automatic cutoff switch, which will run the refrigerator with mains supply when batteries are not charged and when the batteries get charged, it will automatically cutoff the mains supply, thus saving the power consumptions also. Below are the calculations of the power consumed by the SIMR.



Fig. 3 SIMR prototype

Fig. 4 Proposed app



5.1 Power Consumption = 36 W (Peltier Based)

Considering the sunlight available from 11 am to 4 pm (peak hours in Dehradun, India) = 5 h. The battery requirement will follow as:

The SIMR will run on battery for 24-5 (in Hours) = 19 h.

Once precool period is over, fridge runs for 2 h in 3 h and stays switched off for 1 h. That is, the effective h are $19 \times 2/3 = 12.67$ h ≈ 13 h. The energy consumed is $13 \times 36 = 468$ Wh. So, the battery size is Battery = $468/13 = 36 \approx 40$ Ah.

Also, the solar photovoltaic panel requirements are as follows:

The panel has to charge 40 Ah battery in 5 h and side-by-side supply 3 A to the refrigerator, i.e., $40/5 + 3 = 11$ A current is required. Therefore, the panel ratings are $11 \times 18 = 198 \approx 200$ W.

5.2 Power Consumption 140 W (Compressor Based 190 L Capacity)

A modern 140 W unit uses on an average about 1.1 kWh daily, i.e., the compressor stays on for 8 h in 24 h a day. That is, 1/3 time. The sunlight period is from 11 am to 4 pm, i.e., for 5 h. Therefore, the period for which it will run on the battery is 19 h. Thus, the effective time = 1/3 of 19 = 6.333 h and $7\text{ h} + 1/3\text{ of } 5\text{ h} = 8$ h. Now, the battery required is $8 \times 140 = 1100$ W, i.e., $1100/13 = 84$ Ah ≈ 90 Ah so battery required is of 90 Ah.

Also, the panel requirement is as follows:

Time the panel has to charge the 90 Ah battery = $5 - 1/3$ of (5) = 3.33 h and current required is $90/3.33 = 27$ A so the panel rating is $27 \times 18 = 486 \approx 500$ W.

6 Conclusion

The paper is on the designing of the solar-assisted cloud-enabled intelligent refrigerator. The focus of the paper is on the utilization of the wasteover food and to avoid food wastage. The designed system is intelligent and IoT-based operated by an app that automatically notify the food van agencies, NGOs take the food leftover in hotels, societies, or wedding spots, and then further, it will be distributed to the needy people. Furthermore, in future scope, more work can also be done to identify the type of food present in the SIMR using image processing or any other technology.

References

1. McKenna, J.: Pope Francis says wasting food is like stealing from the poor. In: The Telegraph 2013, Telegraph Media Group (2013)
2. Riches, G.: First World Hunger: Food Security and Welfare Politics. Springer, Berlin (2016)
3. India, F.S.F. Hunger in India (2018). <https://www.indiafoodbanking.org/hunger>
4. Mondal, S., et al.: An overview of cleaning and prevention processes for enhancing efficiency of solar photovoltaic panels. Curr. Sci. (00113891) **115**(6) (2018)
5. Mukherjee, M., et al.: A vision of IoT: applications, challenges, and opportunities with Dehradun perspective. In: Proceeding of International Conference on Intelligent Communication, Control and Devices. Springer (2017)
6. robu.in. *Load Cell*. Weighing load cell sensor 5 kg YZC-131 with wires. <https://robu.in/product/weighing-load-cell-sensor-5kg-electronic-kitchen-scale-yzc-131-wires/>. Accessed 27 Oct 2018

Internet of Things in Home Automation—A Review



Aayushi, A. K. Nida Susan, Natasha Sanjay Ayare, N. S. Prema and S. B. Chandini

Abstract Home technology is used to access Android smartphone control devices in your home. The aim of home automation is to improve the ease of control of the house. The task performs automatically and is more comfortable. The various tools used for home automation include Near-Field Communication, Raspberry Pi, Arduino Board, Actuators, Android Devices, etc. The paper discussed about how other researchers used the role of the Internet of Things in home automation in their work.

Keywords Home automation · IoT · Sensor · Smart home · Home appliances

1 Introduction

The Internet of Things (IoT) is a technique of connecting the objects to the sensors collecting information and communicating with the nearby people using wireless, sensor, and mobile technologies. IoTs mainly helps in managing and controlling objects around us in a smarter and meaningful way and also in leading cost-effective, quality, and secured life. With increased technological progress, the dependence of the people has increased. It helps to improve living standards and a safe environment.

A use of smart home systems helps in easy monitoring and controlling of the objects inside the home and provides security system using locking systems. Mobile devices are used to connect users and home appliances. The IoT system includes the following separate components: Devices and Sensors, Data processing unit, Connectivity, and User interface. Any object that contains the computer power and Internet connectivity has the ability to collect and transfer data through a network with-

Aayushi · A. K. Nida Susan (✉) · N. S. Ayare · N. S. Prema · S. B. Chandini
Department of Information Science and Engineering, Vidyavardhaka College of Engineering,
Mysuru, Karnataka, India
e-mail: nidasusan1210@gmail.com

out manual support or interaction. In this paper, we have discussed various home automation systems proposed by different authors and method/devices used were also discussed.

2 IoT in Smart Home Automation

In [1], the authors proposed a system that works with temperature and approximation values. This system consists of three main components, including Raspberry Pi, a Global Positioning System (GPS) mobile device and a cloud-based virtual machine. With the help of a temperature sensor, the Raspberry Pi 3 system is designed to monitor the external temperature of a mechanical actuator. It can send temperature data to the cloud in the database. The mobile device is equipped with an Android application based on GPS, which can send a position report to the cloud in a database. The latest position and temperature data are stored in the database, which is then compared to a predetermined development threshold and the decision is taken to turn the actuator on or off. At each development interval, the system stores and compares cloud data. All processing of data is carried out in the cloud as shown in Fig. 1.

In [2], the authors described the rapid growth of usage and confidence in the dynamic properties of intelligent devices. There is much home automation in existence, but not cost-effective. They offer an efficient home automation system using Wi-Fi. It is based primarily on the concept of device Internet work. An Android application can control these devices. The internet controls this system. It uses sensors to obtain moisture and temperature readings. Not only is the system based on a smartphone, it is also an open-source solution that helps reduce costs as shown in Fig. 2.

In [3], various researchers have studied remote control of household appliances using various technologies. They have proposed an algorithm for storage and retrieval in the distributed environment using web services. They have two basic models, one is a circulating token algorithm, and the other is a knowledge-based resource allocation protocol. A hardware connected to a Raspberry Pi board is developed. They used

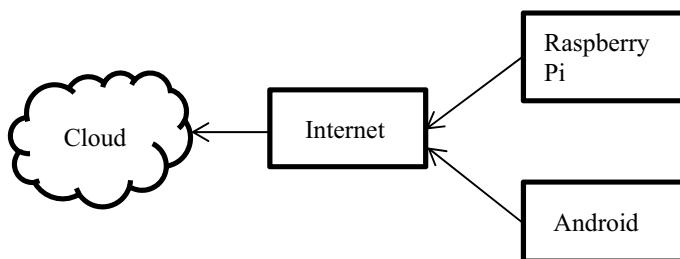


Fig. 1 IoT-based actuator control via cloud for approximation and temperature

native Android languages to create an application. The system's basic operation involves identifying the sensor ID of a specific device and sending the data to the scenario as shown in Fig. 3.

In [4], the authors describe how to control and monitor household appliances with Android apps over the internet. Many home automation systems are available on the market, but they are of limited use. The home appliances can be individually controlled from home and remotely, helping people with physical disabilities. This work was designed to create a virtual but practical home automation system based on Android, as shown in Fig. 4. The Android phone uses Arduino controls to control all appliances in your home. It ensures safety and cost-effectiveness.

In [5], the authors proposed a highly tenacious and environmentally sustainable home automation system, which uses IoT to control home appliances using an Internet access device. The main part of this system is a small Microprocessor, Raspberry Pi, and a Microcontroller, Arduino Uno, and an Android application to view the data supplied by Raspberry Pi and to send, receive, and process requests as shown in Fig. 5. The main objective of this model is to develop a home automation system that interacts with the user with different push notifications based on relevant parameters while respecting the environment. The Home Automation System (HAS) can serve a number of functions, including air control, garden maintenance, harvesting of rain-water, etc. The overall system is low cost and can be improved with functions like voice recognition, facial recognition, and a heartbeat sensor to monitor the health of the user.

In [6] with technology advancement today, the number of household electrical appliances are increasing, particularly for the IoT Smart Home. The author proposed a mobile application that could run on Android-based smartphones to provide high security by implementing a dual encryption method that encrypts the communication data using both RSA and AES algorithms. It has additional security features including permissions and automatic locking. Furthermore, the appliance can con-

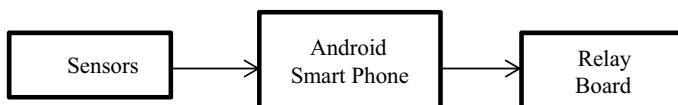
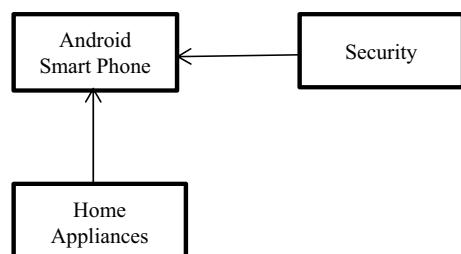


Fig. 2 Low-cost IoT-based home automation system

Fig. 3 Smartphone using IoT security for home appliances



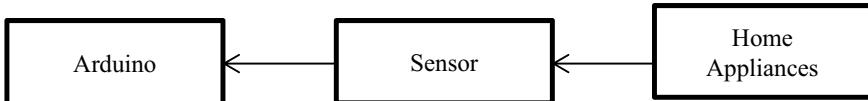


Fig. 4 Arduino-based home appliances control using smartphone

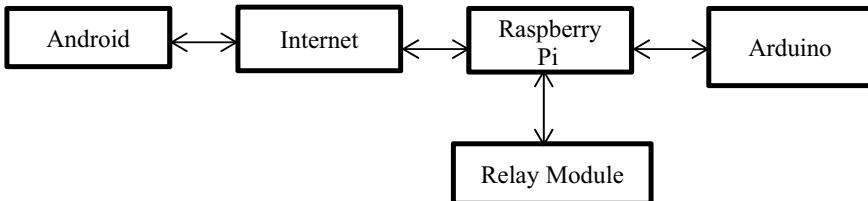


Fig. 5 Enhanced home automation system using IoT

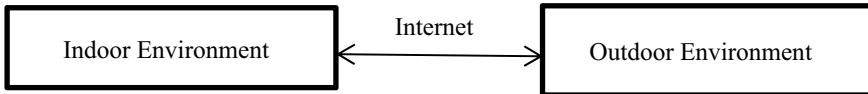


Fig. 6 High security and automatic features with smart home mobile application

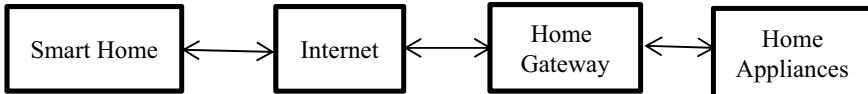


Fig. 7 Home automation via Internet using smartphone

trol the appliances automatically and intelligently using the scenario and planning functionality as shown in Fig. 6.

In [7], the application uses an Arduino-integrated micro-web server with real IP connectivity for Android-based access and control devices. The architecture proposed is composed of three levels—home environment, home gateway, and remote environment. The home environment is home router and interface module. The home router (gateway) function is provided with data translation services between the internet by sending and receiving answers as shown in Fig. 7. It allows authorized homeowners to control and monitor connected home devices remotely via WLAN or 3G/4G that supports Java. Devices such as light switches, sirens, and heat, Humidity, Current, and Smoke sensors have been incorporated to demonstrate the feasibility and efficiency of the developed home automation system.

In [8], a user sensor, Bluetooth, and Android, a home automation system with custom artifacts is proposed. Near-Field Communication (NFC) is one of the most advanced technologies in different fields. NFC is a type of innovation that incorporates a range of conventions, which correspondingly increases the radio repetition between the two gadgets within a short distance of 10 cm or less. The operating fre-

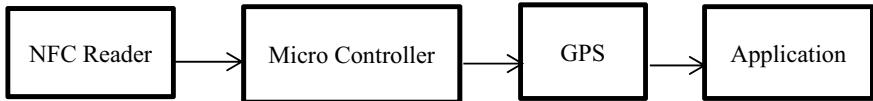


Fig. 8 Home automation using NFC

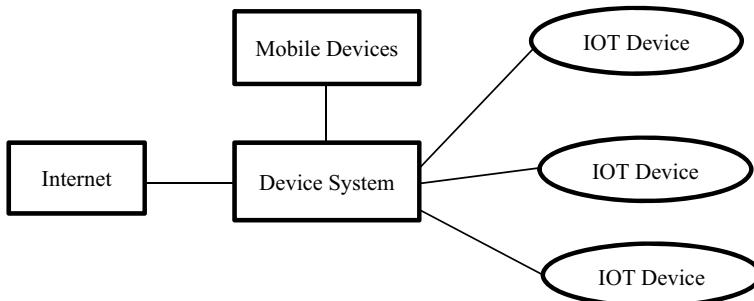


Fig. 9 Secured IoT for smart home system

quency is 13.56 MHz and the exchange rate is 424 kilobits per second. This type of communication helps to automate and remotely control household appliances. The author proposed the home automation architecture using NFC and General Packet Radio Service together with a mobile application. The architecture projected automates and controls the home environment using NFC technology and GPRS monitoring as shown in Fig. 8.

In [9], the author implements the intelligent home system based on IoT Wi-Fi. The proposed architecture uses a home gateway as a center node enabling secure communication between IoT device numbers are shown in Fig. 9. Through mobile device, the users will access, control, and configure the system interface such as Ubiquitous Smartphones. IPv6 allows the connection of a number of embedded devices much more than Zigbee or Bluetooth. This approach is carried out through a public key mutual authentication protocol that uses elliptic curve cryptography because it is highly secured since the use of pre-shared keys eliminates the need for additional public key infrastructure.

In [10], the authors present the design of a smart home automation system that uses the cloud and wireless communication mix to remotely control various electronic devices via a smartphone for users in their homes. The system's architecture comprises of two stations, one base station and one satellite station. The base station is only an Arduino mega-microcontroller board connected to Wi-Fi and thus accessible to the internet as shown in Fig. 10. It consists of a module for the RF transceiver that communicates with the satellite station. The Arduino Uno satellite stations are microcontroller boards with sensors such as light detector, movement sensor, temperature sensor, etc. The Wi-Fi access point is initially connected via the Wi-Fi module when the system is at the base station and the TCP connection is set

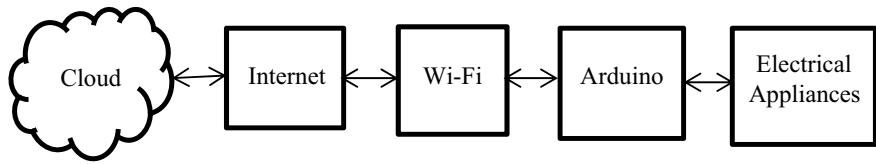


Fig. 10 Customary homes to smart homes using IoT

up. The sensor data and space conditions are collected individually by the base and satellite stations via an RF link when the user sends the command and the data is transferred via the Wi-Fi module to the cloud via the given API. The system automatically controls sensor-based devices by monitoring the home environment and storing sensor data in the cloud continuously.

3 Conclusion

In recent decades, the home automation system has become very popular by making comfortable and quality life. In this paper, the emerging automation systems are discussed. A smartphone is used in controlling and monitoring home appliances using various communications techniques. The main concept of all these papers is to have a low cost and flexible control system based on Android. All systems in their designs are feasible and effective. They are a low-cost system that provides security and automatic control. Many automation systems are proposed, which many people can take usage of it and live in an intelligent environment.

Acknowledgements The authors express gratitude toward the assistance provided by Accendere Knowledge Management Services Pvt. Ltd. In preparing the manuscripts, we also thank our mentors and faculty members who guided us throughout the research and helped us in achieving the desired results.

References

1. Pablo, P., Andres, C.: Approximation and temperature control system via an actuator and a cloud: an application based on the IoT for smart houses. In: 2018 5th International Conference on eDemocracy and eGovernance, ICEDEG 2018, pp. 241–245 (2018)
2. Vikram, N., Harish, K.S., Nihaal, M.S., Umesh, R., Shetty, A., Kumar, A.: A low cost home automation system using wi-fi based wireless sensor network incorporating internet of things (IoT). In: Proceedings - 7th IEEE International Advanced Computing Conference, IACC 2017, pp. 174–178 (2017)
3. Thiagarajan, M., Chaitanya, R.: Integration in the physical world in IoT using android mobile application. In: Proceedings of the 2015 International Conference on Green Computing and Internet of Things, ICGCIoT 2015, pp. 790–795 (2015)

4. Siva Nagendra Reddy, P., Tharun Kumar Reddy, K., Ajay Kumar Reddy, P., Kodanda Ramaiah, G.N., Nanda Kishor, S.: An IoT based home automation using android application. In: International Conference on Signal Processing, Communication, Power and Embedded System, SCOPES 2016 – Proceedings, pp. 285–290 (2016)
5. Sri Harsha, S.L.S., Chakrapani Reddy, S., Prince Mary, S.: Enhanced home automation system using internet of things. In: International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC 2017), pp. 89–93 (2017)
6. Trio, A., Suksmandhira, H., Billy Austen, M., Waskita, A.: Design of smart home mobile application with high security and automatic features. In: IGBSG 2018 - 2018 International Conference on Intelligent Green Building and Smart Grid, pp. 1–4. IEEE (2018)
7. Amritpal, K., Saud, A.: Home automation system via internet using android. *Int. J. Adv. Comput. Eng. Netw.* **4**, 15–18 (2016)
8. Vagdevi, P., Divya, N., Golla Vara, P.: Home: IOT based home automation using NFC. In: Proceedings of the International Conference on IoT in Social, Mobile, Analytics and Cloud, I-SMAC 2017, pp. 861–865 (2017)
9. Santoso, F.K., Vun, N.C.H.: Securing IoT for smart home system. In: Proceedings of the International Symposium on Consumer Electronics, ISCE, pp. 5–6 (2015)
10. Vignesh, G., Mithileysh, S.: Customary homes to smart homes using internet of things (IoT). In: 2017 International Conference on Smart Technology for Smart Nation, pp. 1059–1063 (2017)

A Study on Different Types of Web Crawlers



P. G. Chaitra, V. Deepthi, K. P. Vidyashree and S. Rajini

Abstract The world wide web is a global information medium in which as many people as possible explore the information around the world. Search engine is a place where internet users search for the required content and the results are returned to users through websites, images or videos. Here web crawlers emerged that browses the web to gather and download pages relevant to user topics and store them in a large repository that makes the search engine more efficient. These web crawlers are becoming more important and growing daily. This paper presents the various web crawler types and their architectures. Comparisons are analyzed between these crawlers.

Keywords Web crawler · Focused crawler · Incremental crawler · Distributed crawler · Parallel crawler · Hidden web crawler

1 Introduction

The world wide web has increased rapidly. Millions of people rely on world wide web to seek information each and every second. Everyday the content of web pages gets updated. It is not possible for the search engine to search entire content of the web to retrieve specific web pages requested by the user. Hence search engine depends on web crawlers which crawl the relevant web pages from the world wide web. World Wide Web is a graphic structure where web pages' act as nodes and links. The web crawler passes through the graph by visiting the web pages of a Uniform Resource Locator (URL) seed and moving from one page to another by following the links on the pages. Web crawlers therefore retrieve websites and store them in a

P. G. Chaitra (✉) · V. Deepthi · K. P. Vidyashree · S. Rajini
Department of Information Science and Engineering, Vidyavardhaka College of Engineering,
Mysuru, Karnataka, India
e-mail: chaitrasharma3@gmail.com

V. Deepthi
e-mail: deepthiv029@gmail.com

local repository later used by the search engine that indexes downloaded pages [1]. Automated maintenance of web pages like validating HTML codes and checking links is done by web crawlers. Web crawler keeps track of the web regularly so the results they generate are updated and timely. The nature of the web crawler depends on the following policies. Selection policy states to download the web pages. Revisit policy states to revisit the web pages whenever its content updates. The overloading of websites is prevented by politeness policy. How to synchronize distributed web crawlers is specified by parallelization policy [2].

1.1 Web Crawlers

It consists of frontier, downloader page, web repository as its components. Frontier reserves the list of URLs to visit. The seed URL may be obtained from the user initially. Crawler's work begins from the border-reserved seed URL. The pages that correspond to the seed URL are taken from the web and the list of unvisited URLs from the page is stored at the frontier. Page downloader downloads web pages from the web, corresponding to the frontier URL. Web repository is where web pages are stored and managed. It reserves HTML pages only. It reserves the pages as separate files, and the storage manager keeps all pages accessed by the crawlers updated as shown in Fig. 1.

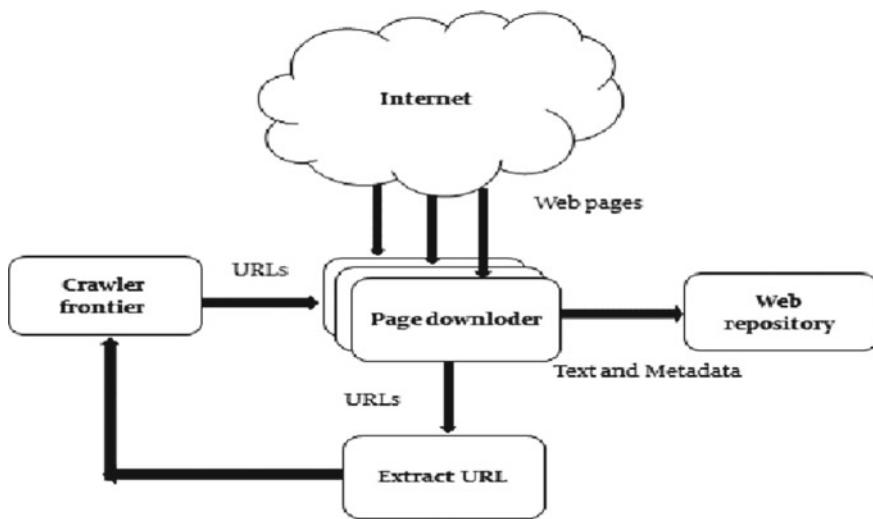


Fig. 1 Architecture of web crawler [1]

2 Types of Web Crawler

2.1 Focused Web Crawler

Focused web crawler selectively search for web pages relevant to specific user fields or topics. It attempts to obtain more relevant pages with a higher level of accuracy. It only downloads relevant pages and avoids pages not related to the subject. This is achieved by prioritizing websites. “Predictor segment” [3] is a technique used that predicts the relevance of the website segment. The Website segment is a web page group found in the same directory path. It concludes the probability that a link is relevant to a particular page before downloading from the web, this can be done by an approach “anchor text of the links” [4]. Page relevance may also be concluded after downloading the websites. “Fish search” is a crawler example that assigns priority values (0 and 1) to pages using “keyword matching”. The disadvantage of using this is that it assigns value 1 to all pages matching the keyword, improved by using a new “shark search” technique. There are different approaches used in order to fetch the relevant pages from the web such as priority-based crawler, structured based crawler, learning-based crawler, and context-based focused crawling [4].

Figure 2 represents the focused web crawler architecture. Initially, the user generates seed URL. These URLs are stored in the web’s “inner Crawler Presentation” [5]. The web graph $G = (V, E)$ consists of vertices and edges corresponding to web pages and hyperlinks. V as a set of nodes includes both visited and unvisited URLs. The URL selector component selects the URL based on “proximity measures”.

The URL selector locates and passes the corresponding URL to the browsing component. The URL can be ignored from the same website. Browsing components

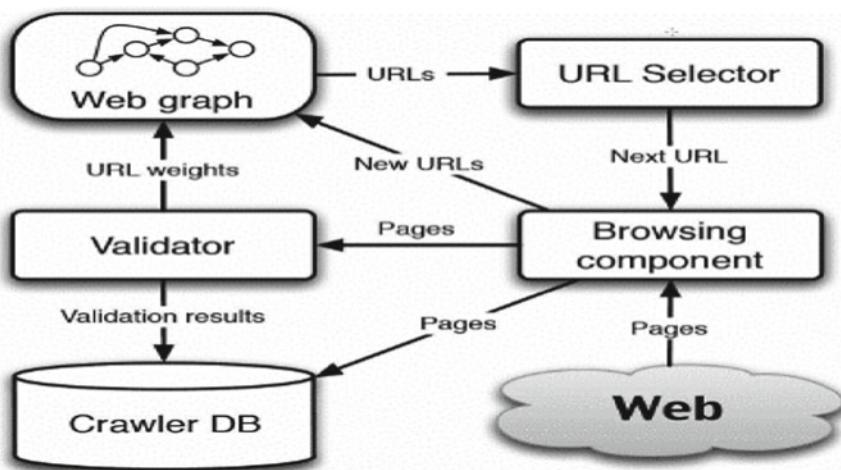


Fig. 2 Architecture of focused web crawler [5]

download the corresponding URL page, process the page, and cross the hyperlinks found on page [5]. The Validator component determines if the page is relevant to a given topic. The validation results, i.e., whether or not the page is relevant, are stored in the database together with the web pages. Various validation techniques depending on the application domain are selected. Some systems use “support vector machines” and “Bayesian networks” [6].

2.2 Incremental Web Crawler

Web crawlers known as incremental crawlers are designed to visit and access updated web pages. Incremental crawlers update the content of websites by visiting them frequently and storing the updated version of pages.

Figure 3 of the incremental crawler is divided into three modules, [7] namely the Ranking Module, the Update Module, and the Crawl Module. The data structure used to implement Incremental Crawler’s work is the priority queue. The priority queue called AllURL stores all URLs discovered by the crawler and CoURL records the URL visited by the crawler. Ranking Module’s task is to select the URL to add to the CoURL. When page not present in the CoURL becomes more important than page present in the CoURL, the Ranking Module is responsible for replacing the page. The URL of the page to be replaced is given the highest priority so that the Update Module can access the page more quickly. Update Module’s main work is to check whether or not web page content is updated. Update Module uses the estimated frequency of change page as the ruling function. The URL is placed in the queue based on these calculations. The URL which is present nearer to the head of the queue will be visited frequently and hence the incremental crawlers provide updated pages to the users.

2.3 Distributed Web Crawler

Distributed crawlers assign crawling to other crawlers. A central server in remote areas communicates and syncs with the nodes. It implements PageRank to enhance its efficiency and quality search [9].

There are two architectures for the distributed web crawling system, namely Master slave and Peer to Peer architecture. The architecture widely used is Master Slave as shown in Fig. 4. It is designed by combining the architecture Scrapy and the Redis database. Master is responsible for scheduled requests and communicates with slaves. Slaves are an independent crawler. The slaves retrieve from the master URL, download the corresponding web pages from the World Wide Web, and return the pages requested to the master. Multiple web pages can therefore be found simultaneously by distributed crawlers.

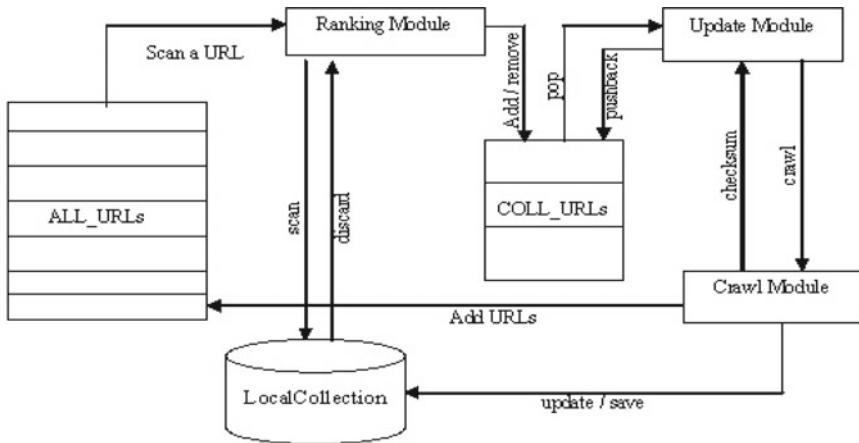


Fig. 3 Architecture of incremental web crawler [8]

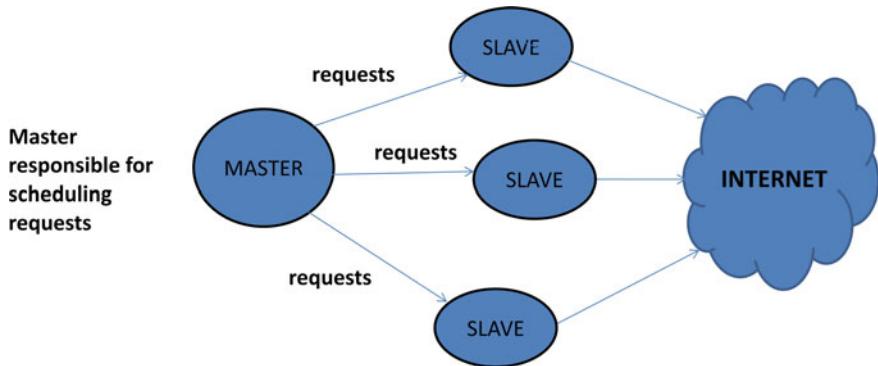


Fig. 4 Architecture of distributed web crawler [9]

2.4 Parallel Web Crawler

The web is growing tremendously and therefore the whole web cannot be crawled through a single mechanism. Therefore, multiple mechanisms of the parallel crawler are created and used that operate parallel [1].

Figure 5 represents Parallel Crawler's architecture. Multiple Parallel Crawler processes are represented as C-proc's. Each C-proc performs a single crawler process task. It retrieves pages from the world wide web, stores pages locally, filters the URLs of the pages retrieved, and passes through these URLs. Collected URLs are sent to the other C-procs by the C-procs. These C-procs can be located on the same local network or in remote areas. On this basis, two architectures for parallel crawlers, such as Intra site parallel crawler and Distributed crawler [1], are available.

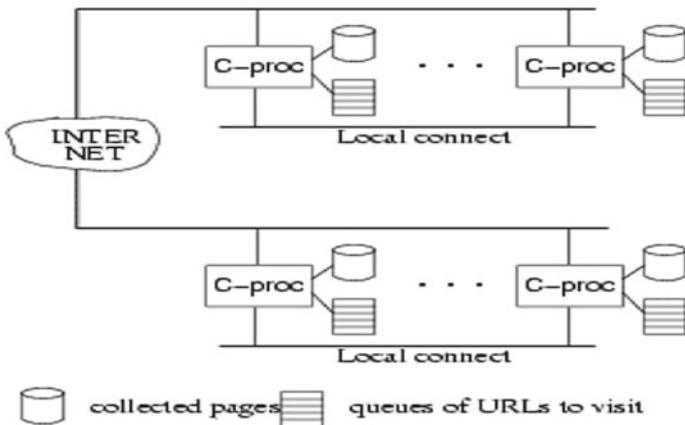


Fig. 5 Architecture of parallel web crawler [1]

Intra site parallel crawler: All C-procs located on the same network interact with each other via a high-speed network.

Distributed crawler: In this architecture, all C-proc's located in far areas interact on the internet.

2.5 Hidden Web Crawler

The world web is divided into “surface web” and “hidden web” parts. Web content that can easily be accessed by general purpose crawlers is referred to as surface web. Huge web content is hidden behind search forms that are not accessible by any standard search engine, these hidden pages are referred to as the hidden web or the deep web. These are only available when the user provides a set of keywords or queries also known as query interfaces. It has two categories [10] such as the content of public access and private access. E-commerce is an example of public access content that includes certain data and products in private databases and is accessible via the Internet. Universities and government institutions are examples of private access content and maintain a private database that cannot be accessed over the Internet, i.e., restricts members or subscribers access to the database. Hidden web crawlers are looking for the search form in each web page visited and this form is automatically filled by the LVS manager and submitted to the web. While the surface web ignores this form, which includes permission or registration.

Figure 6 shows the Hidden Web crawler architecture. It consists of three main components Form analyzer, Form and Answer analyzer [11]. The URL list consists of URL seed generated by crawlers. Crawler Manager controls the unified crawling process that takes the URL from the URL list and downloads the URL pages. For-

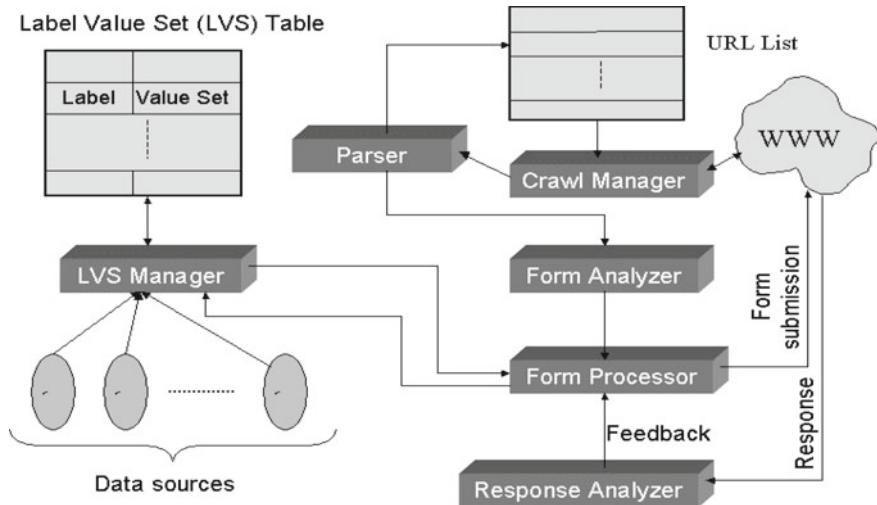


Fig. 6 Architecture of hidden web crawler [11]

wards downloaded pages to the parser that displays and finds hypertext links and adds them to the URL list [7].

The form analyzer parses and process the form when search forms are found on websites. The LVS table values are assigned to the different form elements and submitted to the web by the form processor. Web response pages are analyzed by the answer analyzer and this information is used as feedback for the next iteration. When the links on the response pages are found, they are added to the URL list and visited recursively.

3 Comparison of Different Types of Web Crawler

The different types of web crawlers are compared with the parameters to discussed as shown in Table 1.

4 Conclusion

This paper discussed about various types of crawlers. Each crawler is specific to the specific application. For example, incremental crawlers are used to retrieve updated pages such as news, Hidden crawlers crawl the hidden content of search forms. Each crawler is analyzed in this paper and compared to different parameters.

Table 1 Comparison of different types of web crawler

Sl. no.	Parameters	Hidden crawler	Distributed crawler	Incremental crawler	Parallel crawler	Focused crawler
1	Freshness	No	No	Yes	No	No
2	Search technique	DFS	BFS	BFS	BFS	DFS
3	Network load reduction	–	No	No	Yes	–
4	Scalability	Yes	No	No	Yes	Yes
5	Extensibility	–	No	Yes	No	–
6	Overlapping	–	No	No	Yes	–
7	Selection of pages	Form analyzer	From seed URL's	From priority queue	From seed URL's	Related to specific topic

Acknowledgements The authors express gratitude towards the assistance provided by Accendere Knowledge Management Services Pvt. Ltd. In preparing the manuscripts. We also thank our mentors and faculty members who guided us throughout the research and helped us in achieving desired results.

References

1. Gupta, S.B.: The issues and challenges with the web crawlers. *Int. J. Inf. Technol. Syst.* **1**, 1–10 (2012)
2. Castillo, C.: Effective web crawling. Ph.D. thesis. University of Chile (2004). Accessed 03 Oct 2018
3. Suechua, T., Rungsawang, A., Yamana, H.: Adaptive focused website segment crawler. In: 19th International Conference on Network-Based Information Systems, pp. 181–187 (2016)
4. Gupta, A., Anand, P.: Focused web crawlers and its approaches. In: 2015 1st International Conference on Futuristic Trends on Computational Analysis and Knowledge Management ABLAZE 2015, pp. 619–622 (2015)
5. Shchekotykhin, K., Jannach, D., Friedrich, G.: xCrawl: a high-recall crawling method for web mining. *Knowl. Inf. Syst.* **25**, 303–326 (2010)
6. Yu, H., Han, J.: PEBL: positive example based learning for web page classification using SVM. In: Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2002)
7. Sharma, S., Gupta, P.: The anatomy of web crawlers. In: International Conference on Computing, Communication and Automation ICCCA 2015, pp. 849–853 (2015)
8. Hall, W., De Roure, D., Shadbolt, N.: The evolution of the web and implications for eResearch. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **367**, 991–1001 (2009)
9. Yuhao, F.: Design and implementation of distributed crawler system based on Scrapy. In: IOP Conference Series: Earth and Environmental Science, pp. 1–5 (2018)

10. Kumar, D., Mishra, R.: Deep web performance enhance on search engine. In: International Conference on Soft Computing Techniques and Implementations, ICSCTI 2015, pp. 137–140 (2015)
11. Raghavan, S., Garcia-Molina, H.: Crawling the hidden web. In: 27th VLDB Conference, Roma, Italy, pp. 1–10 (2010)

Security and Trust in IoT Communications: Role and Impact



Keshav Kaushik and Kamalpreet Singh

Abstract With the advent of new technologies, many computational devices with sensing and networking power are connected with each other. These devices sense the data from the environment, perform some computation on it, and then either store or transmit the data to some other device having similar capabilities. With this process of communication, there arise several security and trust concerns, which are very critical to handle. Such security and trust concerns can be potential threats to the IoT ecosystem and they need to be handled with the latest architecture, framework, and technology. This paper attempts to highlight some major threats of security and trust in IoT and how existing architecture deals with them.

Keywords IoT · Security · Trust · TRE · Verification and Validation

1 Introduction to IoT Communications

In modern era, where multiple physical things are connected with each other with internet [1], they sense, compute, and communicate. In order to make Internet of Things (IoT) fully functional in a smooth manner, communication among various things plays a crucial role [2]. Actual communication in IoT is between computational devices, mobile actuators, smart sensors, and embedded processors with or without intervention of humans. When any thing having sensing and networking power is connected at any place and that connection is available any time then it follows the principles of IoT. IoT communications can be applied in various fields such as Health care, smart homes, food nutrition, smart cities, e-business and retail, transportation, etc. With this wide spectrum of IoT communications, there arises the security, privacy, and trust concerns, which needs to be addressed in an efficient and

K. Kaushik (✉) · K. Singh
School of Computer Science, UPES, Dehradun, India
e-mail: keshavkaushik@ddn.upes.ac.in

K. Singh
e-mail: kpsingh@ddn.upes.ac.in

smart way. As connected devices are increasing day by day which is resulting in increase of security threats in IoT, verification and validation of trustworthiness of any IoT device needs to be done with strict compliance [3]. This paper is divided into five major sections. The first section gives a brief introduction about IoT communication and its applications. The requirement for security in various IoT applications is discussed in second section. The third section explains the major categorization of security threats in IoT. The fourth sections give a brief overview about trust-based environment in IoT with a generic Trust based architecture and fifth section gives overview of verification and validation process in IoT. Lastly, it concludes with the findings from this research.

2 Requirement for Security in Various IoT Applications

IoT has its wide spectrum of applications, be it smart cities, healthcare, transportation, wearables, smart homes, e-business and retail, smart grids, etc. Now with this wide range of applications, threats and vulnerabilities are growing day by day, which makes securing IoT as a need of the hour. This section covers the need and requirement of security and trustworthiness in various applications of IoT. Communication between billions of connected devices is making implementation of security in IoT of supreme importance. Different use cases are discussed below:

2.1 Smart Meter Reading

Measuring the usage of house and office utilities like water and electricity is a time and cost consuming process. Smart Meter Reading refers to the technology used for automatic collection of meter readings from various meters and transfer of this information to central database for billing and analysis purpose [4]. With increasing acceptance of these smart technologies, the requirement of real time monitoring of utility usage also increases so as to optimize the use of smart grids.

2.2 Tracking of Consignments

Remote Tracking System allows the buyers/owners of goods to invigilates certain crucial parameters and monitor the movement of equipment by performing remote commands. IoT devices in asset tracking are usually placed in places where manual access of people is difficult. Service provider needs to take care of physical tampering of IoT devices and abide theft. Placement of such IoT devices is a critical process [5] as these devices may also be mobile making it costly and difficult to access these devices physically.

2.3 Vending Machines

These are automated machines that provide an efficient and economical way of distributing goods to customers. Any type of attack on the content of these machines hinders the electronic collection of payments and increases threat to other valuable items in machine. Hence, the service provider needs to keep the track of proper functioning of machine along with the stock levels. In certain scenarios, vending machines may also be desired to display multimedia text/videos for marketing purposes. From the connecting options available within the customer premises, one such option may be used for connecting vending machines.

2.4 Traffic Surveillance

Cameras installed in remote locations. Footbridges and flyovers are used to analyze the traffic and to spot over speeding vehicles. For calculating average speed and proper monitoring, these cameras need to be cellular connected (to at least the next camera on the road) in a secure private WLAN connectivity [6].

3 Security Threats for IoT

Talking about IoT devices, they are small, inexpensive, low cost, ability to operate in an unattended environment for a long period. IoT devices are designed in such a way that they can communicate with each other over WLAN and they can be deployed by humans and can be controlled and managed remotely [7, 8]. In addition, these devices are mobile in nature and are deployed on large scale in many cases. All these requirements of IoT devices give rise to security vulnerabilities in wireless communication of IoT devices. These security threats and vulnerabilities are broadly classified into following categories:-

- Physical Attacks: These types of attacks deal with tampering of hardware components, input of authentication tokens into manipulated devices, booting of devices with malicious code by flashing them on the device memory. Such types of attacks are difficult to perform because of their expensive nature [9]. Some examples of such attacks include Reverse Engineering, Micro Probing attaching microscopic needles onto the internal wiring of a chip, which read out internal secrets, chip de-packaging, reconstruction of hardware layout, particle beam techniques, tags cloning and tampering, hardware tampering [10].
- Configuration attacks: These attacks deal with misconfiguration or improper configuration of hardware devices, which leads to installation, or updating of rogue or fraudulent software [8]. Results of such kind of attacks are very dangerous as it changes the default of normal working of devices by compromising the access control list (ACL).

- Attacks on credentials: These attacks target the weak authentication algorithms, side channel attack, compromising the credentials with brute force method. It also targets Machine Communication Identity Module (MCIM) by fraudulently cloning the authentication tokens [11]. Default passwords of IoT devices are soft targets for such type of attacks.
- IoT protocols attacks: All the attacks targeting the IoT protocols comes under this category. Major examples are Man in the Middle attack (MITM), Denial of Services (DOS), Logic bombs, black hole attacks, selective forward attacks, wormhole attack, Sybil attack, etc. Such types of attacks target the active services on IoT devices by exploiting the vulnerabilities borne due to IoT protocols.
- Core Network attacks: These types of attacks target the Mobile Network Operator (MNO) associated with communications among IoT devices [12]. It works by impersonating the IoT devices by tunneling the traffic between them. Such attacks may change the authorized physical location of a device by spoofing it with some other location; it also targets the gateways/modems/routers by injecting malicious code in them.
- Identity-based user data attacks: Such type of attacks target the privacy of user and users data based on identity [9]. It includes eavesdropping over the network to get the users data and identity or masquerading as other device or subscriber. It also targets users network identity in order to reveal confidential information of a user as an unauthorized third party.

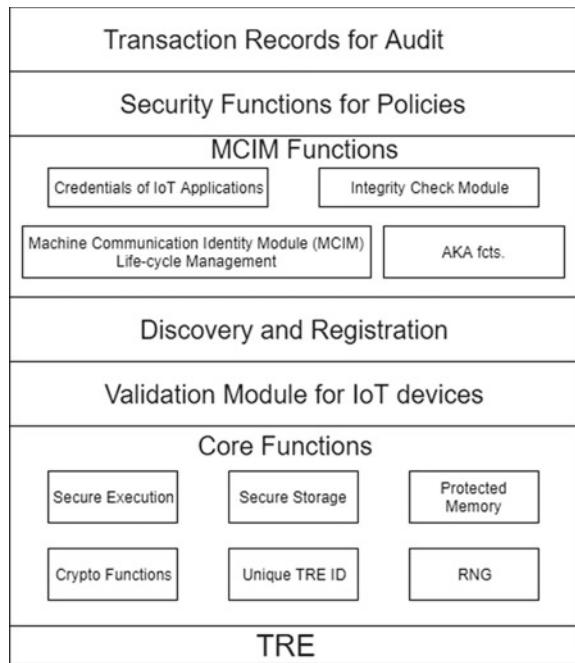
4 Trust Based Environment in IoT

In the current scenario, where millions of connected things are communicating with each other, trust plays a very crucial role. In order to maintain a trust relationship in distributed system, it is important that the system must bear a capability to form the building blocks of trust boundaries. In IoT, a Trusted Environment (TRE) constructs a system, which encapsulates the features of enforcement and trust. TRE is a logical separation of various entities within IoT devices.

TRE actually gives an environment for storage of sensitive data and its execution in a trustworthy manner. It also prevents unauthorized access of user data by separating software and data from the IoT device [7]. Some security and privacy problems are resolved by providing the hardware security measures. Root of Trust (ROT) secures internal system operations based on abstraction. A secure start up process of TRE makes sure that it reaches its trustworthy state. All programs and components that are executed in the secure boot process provide trust to operating system and system software. Whenever a new component is loaded to the system its trust is verified by measuring the integrity at the time of initialization [13].

This trust model uniquely identifies configuration and state of every component. TRE protects the integrity and the measurement results of new components. The process of verification of measurement values with reference values results in acceptance or rejection of new component in trust boundary. As this verification process occurs locally, it assumes that TRE is in its predefined state after completion of execution.

Fig. 1 Architecture of a trusted environment (TRE) in IoT



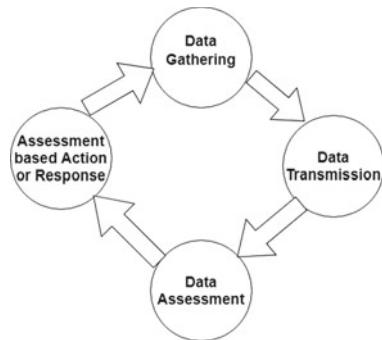
After verification, there comes the process of validation of components, which can be done by transferring the result of verification to some external party [14]. Based on values of TREs functions, validation module predicts the trustworthiness. After performing the trustworthiness of IoT device [15], TRE gives some functions for authentication of IoT devices to the network of operator. The authentication data is stored inside the TRE in order to confirm the authentication of any IoT device to network.

Figure 1 shows the architecture of a Trusted Environment (TRE) in IoT, which comprises of multiple modules like- Transaction records for audit, security functions for policies, Machine Communication Identity Module (MCIM) functions which is responsible for integrity check, registration and discovery module, some core functions and validation module for IoT devices. All these modules combined to form a trusted environment for IoT.

5 Verification and Validation of Trust

There are two main security aspects in machine to machine application scenario
 (a) The unexpected connection of the equipment to the network. (b) The connected equipment needs to be highly flexible and configurable. The main goal of security is IoT communication is to:

Fig. 2 Validation process in IoT communications



1. ensure reachability and operability of M2M device locally in a secure manner without being connected with the core network.
2. facilitate the formation of trustworthy operations on M2M devices to enhance its security properties both globally and locally.

Hence, these protection goals play a vital role when M2M devices change their state (in a controlled way) either being stationary or mobile. Validation refers to the process of technically assessing the system for security-related parameters, (i.e., of verification process). Validation encompasses over the entire lifecycle of an IoT device starting from the initial phase of development (to validate installation verification) to maintenance (to validate the success). Validation is categorized in the following three categories:

- Self-Governing Validation: This validation is applicable to a closed system that do not communicate with the exterior world, e.g., smart cards. Security in such cases is to be handled locally. In this type of validations, prior to communication with the outer world, IoT device is verified locally. Validation process is independent of any external entity. Since the verification process is not shared with the external world, this local verification is assumed to be more secure [16]. Any device failing this verification process is debarred from being connected to the network. Trust model used for smart cards is an example of such validation.
- Isolated Validation: In contrast to the autonomous validation, this validation abstraction of remote attestation.
- Semi-Autonomous Validation: This validation lies between the two extreme ranges of validations discussed above.

Figure 2 shows the validation process occurs throughout the lifecycle starting from data gathering to action/response.

6 Conclusion

The presented paper enables to explore various business opportunities along with the motive of achieving security which is hardware-backed. With the increasing use of IoT applications in various business applications, the need to guard these distributed devices from various security threats have also emerged. Various scenarios for IoT communication which evolve the need of security either due to deployment of IoT devices in the field or their requirement of being flexible. These requirements require a blended mixture of traditional security enforced along with device-oriented trust. Distribution of this mixture of trust building and security enforcement between IoT devices and network help in providing flexibility and scalability to the concepts of technical tasks. Local state control and semi-automatic validation act as the two building blocks for security enforcement. With embracement of these security concepts, the requirement of IoT devices of management of platform remotely and management of subscription can be fulfilled.

References

1. Haldar, V., Chandra, D., Franz, M.: Semantic remote attestation: a virtual machine directed approach to trusted computing. In: USENIX Virtual Machine Research and Technology Symposium (VM 04), pp. 29–41 (2004)
2. Usman, M., Ahmed, I., Aslam, M.I., Khan, S., Sha, U.A.: SIT: a lightweight encryption algorithm for secure internet of things. Int. J. Adv. Comput. Sci. Appl. IJACSA 8(1) (2017)
3. Want, R., Dustdar, S.: Activating the internet of things [guest editors introduction]. Computer **48**(9), 16–20 (2015)
4. Auto-Id Labs. <http://www.autoidlabs.org>. Accessed 4 Sept 2017
5. TCG. Trusted computing group architecture overview version 1.4 (2006)
6. Simplicio, Jr., M.A., Silva, M.V., Alves, R.C., Shibata, T.K.: Lightweight and escrow-less authenticated key agreement for the internet of things. Comput. Commun. (2016)
7. Karygiannis, T., Eyd, B., Barber, G., Bunn, L., Phillips, T.: Guidelines for securing radio frequency identification (RFID) systems. NIST Spec. Publ. **80**, 1–154 (2007)
8. Sakamura, K.: Challenges in the age of ubiquitous computing: a case study of T-engine an open development platform for embedded systems. In: Proceedings of ICSE06, Shanghai, China (2006)
9. The EPCglobal architecture framework, EPCglobal final version 1.3. www.epcglobalinc.org
10. INFSO Networked Enterprise, RFID INFSO G.2 Micro and Nanosystems. In: Co-operation with the Working Group RFID of the ETP EPOSS, Internet of Things in 2020, Roadmap for the Future, Version 1.1, 27 May 2008
11. TCG, TCG Infrastructure Working Group. Architecture part II - integrity management. Specification version 1.0 revision 1.0 (2008)
12. Presser, M., Gluhak, A.: The Internet of Things: connecting the real world with the digital world, EURESCOM mess@ge, The Magazine for Telecom Insiders, vol. 2 (2009). <http://www.eurescom.eu/message>
13. Chaum, D.: Security without identification: transaction systems to make big brother obsolete. Commun. ACM **28**(10), 1030–1044 (1985)
14. Chen, L., Landfermann, R., Lhr, H., Rohe, M., Sadeghi, A.-R., Stble, Ch., Grtz, H.: A protocol for property-based attestation. In: Proceedings of the First ACM Workshop on Scalable Trusted Computing (STC 06), pp. 7–16 (2006)

15. Sadeghi, A.-R., Stble, Ch.: Property-based attestation for computing platforms: caring about properties, not mechanisms. In: Proceedings of the 2004 Workshop on New Security Paradigms NSPW 04, pp. 67–77 (2004)
16. de Laat, C., Gross, G., Gommans, L., Vollbrecht, J., Spence, D.: Generic AAA Architecture IETF Network Working Group. RFC 2903

Optimization of Log Gabor Filters Using Genetic Algorithm for Query by Image Content Systems



N. Jyothi, D. Madhavi and M. R. Patnaik

Abstract Query By Image Content (QBIC) system permits the user to extract similar images from the large database by specifying the query image. In this paper, a new hybrid QBIC system is proposed, which effectively utilizes the color, texture, and shape features by sequential process of three pipe stages which gives good retrieval performance and also reduces computational complexity. In first pipe stage, color features are extracted with histograms based on HSV color space. In second pipe stage, texture information is obtained by optimizing log Gabor filters using genetic algorithm. Finally, in third pipe stage, shape features are extracted with signature functions using polygonal algorithm. Simulation results show that the proposed hybrid sequential pipe stage combination gives good performance metrics in terms of mean recall and mean precision and also reduces the number of filters compared to the existing algorithms.

Keywords Genetic algorithm · Log Gabor filter · Mean precision · Mean recall

1 Introduction

There is a huge demand for retrieval of similar images from the large database. It is a very challenging research area in different applications like Google retrieval, digital multimedia, etc. In literature, many techniques are available to extract tex-

N. Jyothi (✉) · D. Madhavi

Department of EIE, GIT, GITAM Deemed to be University, Visakhapatnam 530045, Andhra Pradesh, India

e-mail: njyothi336@gmail.com

D. Madhavi

e-mail: dmadhavi336@gmail.com

M. R. Patnaik

Department of Instrument Technology, AU College of Engineering, Andhra University, Visakhapatnam 530003, Andhra Pradesh, India

e-mail: ramesh_patnaik@yahoo.com

ture, color, and shape features. Chisti et al. [1] applied optimized Gabor filters using particle swarm optimization and genetic algorithms to identify defective items in automation of industry products. Madhavi et al. [2] applied optimized Gabor filters with genetic algorithm for texture analysis in Content-Based Image Retrieval (CBIR) system design. Texture analysis using Gabor wavelet functions compared with multi-resolution texture features by Manjunath and Ma [3] and proved that Gabor analysis gives good retrieval accuracy. Madhavi et al. [4, 5] used tuned color Gabor filters for texture feature extraction in hybrid CBIR system development. Shrivastava et al. [6] identified that the order to achieve better accuracy of CBIR system is by processing color, texture, and shape, respectively. Tyagi [7] reviewed various latest development methods for CBIR systems. Nava [8] has done experimental evaluation of the performance of Gabor and Log Gabor filters using two distance metrics and confirmed that the Log Gabor filters provide good accuracy for texture extraction as compared to Gabor filters. Arrspide et al. [9] used Log Gabor filters for vehicle identification automatically.

In this paper, a new QBIC system is proposed with hybrid pipe stage processing using Log Gabor wavelet functions which are optimized using genetic algorithm. The proposed retrieval system utilizes the hybrid combination of three pipe stages. The main reason for choosing Log Gabor wavelet functions is to eliminate the limitation of uniform coverage of frequency-domain spectrum due to high DC component of Gabor filters and to optimize the number of filters.

The organization of the paper is as follows. Section 2 presents the proposed QBIC system development. Section 3 describes experimental evaluation and results of the proposed method compared to the existing methods. Conclusions are given in Sect. 4.

2 Proposed Methodology

The proposed functional block diagram of a hybrid three pipe stages retrieval system is shown in Fig. 1. The database which consists of P images is the input of pipe stage 1. From this stage, Q best matched images are identified based on color feature similarity and using histogram plots with HSV color space. These Q images are applied as input data set for second pipe stage. In this stage, a set of Log Gabor wavelet functions are applied to capture texture information and these are tuned using genetic algorithm to optimize the number of wavelet functions. The best R images which are nearest to the query image are obtained. In the third pipe stage, extraction is handled using shape features, and here polygonal fitting algorithm is applied to extract similar type of images. The best S images which are best match with query image are obtained.

The extraction of images in each pipe stage is considered as $P > Q > R > S$.

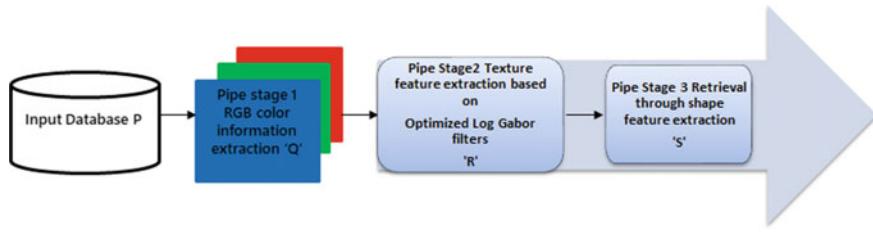


Fig. 1 Proposed pipe stage structure of QBIC system

2.1 *Retrieval of Images Using Color Similarity (Pipe Stage1)*

To retrieve Q images from P input database images, hue, saturation, and value (HSV) color space is used which separates the color information and gray level effectively as compared to RGB color space. The histograms of the P input database images and the query image are evaluated. The similarity is computed using distance metric which is given in Eq. (1).

If $f[0, 1, \dots, n - 1]$ and $g[0, 1, \dots, n - 1]$ are histograms of size “n” then the distance D between these two is measured by Eq. (1)

$$D = \frac{\sum_{i=0}^{n-1} \min(f[i], g[i])}{\min(|f|, |g|)} \quad (1)$$

Here, $|f|$ and $|g|$ are magnitudes of two histograms f and g , respectively.

This distance metric between query image and all input database images is arranged in increasing order and first Q images are considered as best similar images to query image.

2.2 *Retrieval of Images Using Log Gabor Filters with Genetic Algorithm Optimization (Pipe Stage 2)*

The frequency-domain representation of Log Gabor filter function is given by Eq. (2).

$$F(w) = e^{\left(\frac{-(\log(w/w_0))^2}{2(\log(\sigma))^2}\right)} \quad (2)$$

Here, $w_0 = 1/\lambda$, where “ λ ” varies with specific scale.

A tuned Log Gabor filter is designed using GA optimization with scale, orientation, and frequency as parameters for optimization. This filter is referred to as a multi-objective minimization problem. In three different directions, these Log Gabor filters are tuned using genetic algorithm to optimize the fitness function of the energy responses. The energy response of query image and all input images of this pipe stage

2 is computed. The energies of the input images that are nearest to the query image energy response are considered as most similar images to the query image. The first R images which are nearest are considered as best match images.

2.3 Shape Feature Extraction with Polygonal Algorithm (Pipe Stage3)

The best R images that are obtained based on texture feature similarity from pipe stage 2 are applied as input dataset for pipe stage 3. In this stage, the most S similar images to the input query image are extracted by applying signature functions with polygonal algorithm. The Fourier coefficients of these functions describe the shape information which are invariant to rotation, translation, and scaling.

The Fourier domain expression of the signature function $g(t)$ is given by Eq. (3)

$$b_k = \frac{1}{N} \sum_{t=0}^{N-1} g(t) e^{\frac{-j2\pi kt}{N}}; k = 0, 1, 2, \dots, N - 1 \quad (3)$$

The magnitudes of these coefficients $\{b_k\}$ are shape descriptors. The feature vectors of R input images in this pipe stage 3 and the query image are estimated and the similarity is calculated using Euclidean distance metric which is given by Eq. (4)

$$d = \sqrt{\sum_{i=0}^{N-1} |g_q^i - g_d^i|^2} \quad (4)$$

where g_q is query image feature vector and g_d is feature vectors corresponding to R input data images of pipe stage 3. The first S images which are having minimum distance are considered as best matching images to the query image.

3 Experimental Evaluation and Results

The performance of the proposed system is verified using MATLAB with COREL database [10] by taking 1000 images of 10 sets. These sets contain Horses, Beaches, African people, Elephants, Foods, Buildings, Buses, Flowers, Mountains, and Dinosaurs.

The horse image is considered as a query image and is shown in Fig. 2. The values of Q, R, and S are considered as 10, 8, and 4, respectively. The proposed QBIC system results are shown in Fig. 3.



Fig. 2 Query image

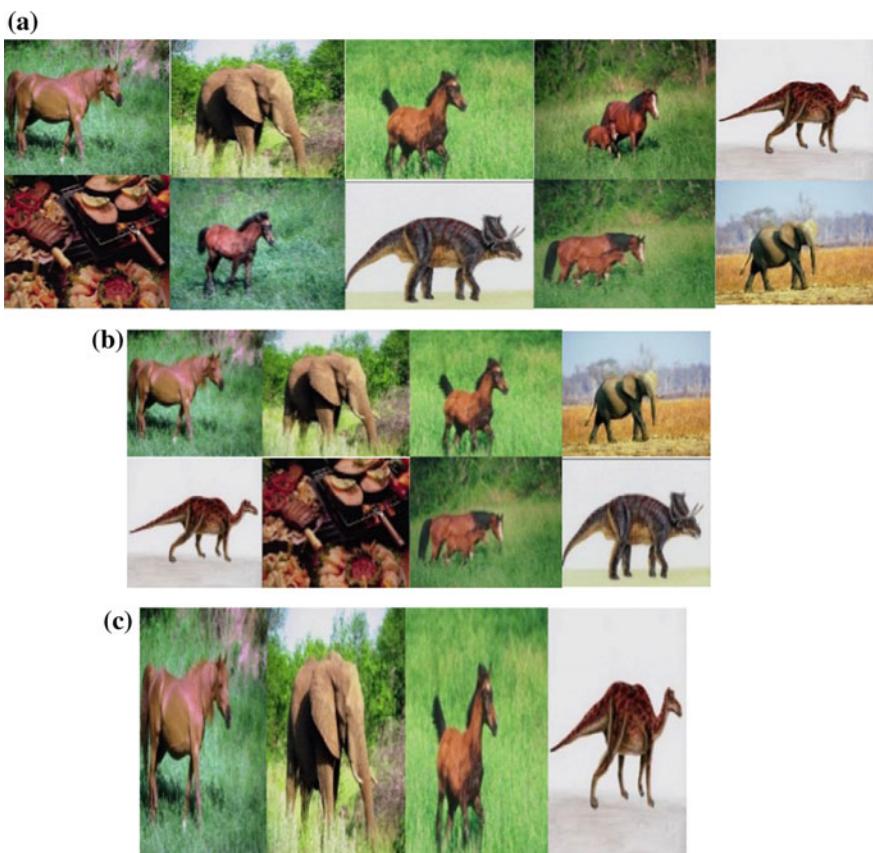


Fig. 3 **a** Computational results of first pipe stage with $Q=10$, **b** computational results of second pipe stage with $R=8$, **c** computational results of third pipe stage with $S=4$

Table 1 Comparison of mean precision of the proposed GA-based optimized Log Gabor QBIC system and other three existing models for 10 classes

Category ID	Class	Log gabor QBIC optimized using GA	Shrivastava model	EIAAlami model	Chuen model
			[12]	[11]	[13]
1	Buses	0.92	0.8	0.87	0.88
2	Food	0.83	0.8	0.74	0.73
3	Horses	0.92	0.9	0.83	0.8
4	Buildings	0.8	0.6	0.57	0.54
5	Elephants	0.92	0.75	0.67	0.65
6	Beaches	0.8	0.6	0.56	0.54
7	African people	0.83	0.75	0.7	0.68
8	Flowers	0.91	0.92	0.91	0.89
9	Dinosaurs	1	1	0.97	0.99
10	Mountains	0.77	0.58	0.53	0.52
Mean		0.87	0.77	0.735	0.722

The performance of the proposed QBIC system is evaluated using standard performance metrics mean precision (MP) and mean recall (MR) which are defined by Eq. (5)

$$Precision = \frac{N_{RI}}{T_I} \quad \& \quad Recall = \frac{N_{RI}}{N_{RD}} \quad (5)$$

Here, N_{RI} is the number of relevant images retrieved, T_I is total images retrieved, and N_{RD} is number of relevant images in database.

Table 1 shows the comparison of experimental values of mean precision of proposed GA optimized Log Gabor filter with three other existing models [11–13] for 10 different classes. It is confirmed that the proposed system shows good mean precision in all the classes of images as compared with other three existing models. Figures 4 and 5 demonstrate the mean precision and mean recall of the proposed system and existing models by varying the number of retrieved images from 20 to 100. The experimental results prove that the proposed system gives better mean precision and recall compared to existing models.

Fig. 4 Comparison of mean precision

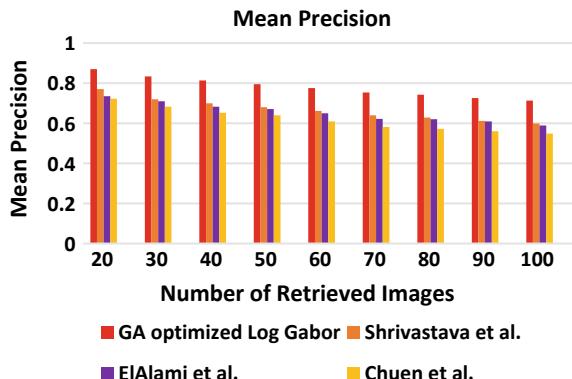
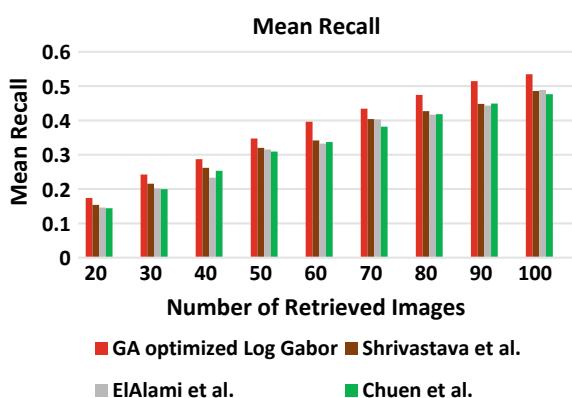


Fig. 5 Comparison of mean recall



4 Conclusion

In this paper, a QBIC system with sequential process of three pipe stages has been proposed. The novelty of the design comes with optimization of Log Gabor filters using genetic algorithm for texture feature extraction which shows better performance metrics in terms of mean precision and mean recall as compared to other existing models.

References

1. Chisti, K.M., Srinivas, K.S., Prasad, G.: 2D Gabor filter for surface defect detection using GA and PSO optimization techniques. *AMSE J.* **58**, 67–83 (2015)
2. Madhavi, D., Patnaik, M.R.: Image retrieval using GA optimized Gabor filter. *Indian J. Sci. Technol.* **9**(44), 1–11 (2016)
3. Manjunath, B., Ma, W.: Texture features for browsing and retrieval of image data. *IEEE Trans. Pattern Anal. Mach. Intell.* **18**(8), 837–842 (1996)

4. Madhavi, D., Patnaik, M.R.: Image retrieval based on tuned color Gabor filter using genetic algorithm. *Int. J. Appl. Eng. Res.* **12**(15), 5031–5039 (2017)
5. Madhavi, D., Patnaik, M.R.: Genetic algorithm based optimized Gabor filters for content based image retrieval. *Advances in Intelligent Systems and Computing*, vol. 624 (2017)
6. Shrivastava, N., Tyagi, V.: An efficient technique for retrieval of color images in large databases. *Computers and Electrical Engineering*. Elsevier, Amsterdam (2014)
7. Tyagi, V.: Content-Based Image Retrieval Techniques: A Review, pp. 29–48. Springer, Singapore (2017)
8. Nava, R., Ramirez, B.E., Cristobal, G.: Texture image retrieval based on Log-Gabor features. *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, vol. 7441. Springer, Berlin, pp. 414–421 (2012)
9. Arrspide, J., Salgado, L.: Log-Gabor filters for image-based vehicle verification. *IEEE Trans. Image Process.* **22**(6), 2286–2295 (2013)
10. Wang Database: <http://wang.ist.psu.edu/>
11. Elalamy, M.E.: A novel image retrieval model based on the most relevant features. *Knowl.-Based Syst.* **24**, 23–32 (2011)
12. Shrivastava, N., Tyagi, V.: An efficient technique for retrieval of color images in large databases. *Comput. Electr. Eng.* **46**(C), 314–327 (2015)
13. Chuen, L.H., Chen, R.T., Chan, Y.K.: A smart content-based image retrieval system based on color and texture feature. *Image and Vision Computing*. **27**, 658–665 (2009)

Self-balancing of a Bike Using Gyroscopic Force and PID Controller



Animesh Jain, Satyam Bhaskar, Kunal Nandanwar and Hari Om Bansal

Abstract The high degree of maneuverability, compactness, and ease, that the modern two-wheelers provide, is undeniably important features in today's domestic environment of road transport. However, the difficulties in balancing these two-wheelers in hilly terrains, the increased uncontrolled traffic with teenagers and carouser driving disastrously, and the accidents caused due to imbalance also need to be noted. Maintaining the pros of the vehicle, a concept of self-balanced bike is proposed in this work. This work proposes a lab prototype of a PID control tuned, self-balancing bike that can be further extended to autonomous bike with path and obstacle detection to travel using global positioning system.

Keywords PID control tuned · Self-balanced bike · Gyroscopic force

1 Introduction

A self-balancing bike is certainly not a new idea. People are working in the field and several theories are there to balance a bike in running as well as standstill position. One of the commercially available products produced by Segway uses two wheels

A. Jain

Department of Mechanical Engineering (B. E. (Hons.)), Department of Biological Science (M. Sc. (Hons.)), Birla Institute of Technology and Science, Pilani campus, Pilani 333031, India

S. Bhaskar

Department of Electrical and Electronics Engineering (B. E. (Hons.)), Department of Physics (M. Sc. (Hons.)), Birla Institute of Technology and Science, Pilani campus, Pilani 333031, India

K. Nandanwar

Department of Mechanical Engineering (B. E. (Hons)), Birla Institute of Technology and Science, Pilani campus, Pilani 333031, India

H. O. Bansal (✉)

Department of Electrical and Electronics Engineering, Birla Institute of Technology and Science, Pilani campus, Pilani 333031, India

e-mail: hobansal@gmail.com

connected perpendicular to the direction of travel, that is, two wheels on a single axle. But this system requires human to continuously interfere with the machine [1]. Similarly, current two-wheelers with internal combustion engine are not designed to balance the structure. Recently, the self-balancing bike developed by Honda boasts to be able to balance the bike at lower speeds without the use of any secondary mass like a flywheel and avoiding gyroscopic principle. Their theory works by changing the trail length from positive to negative as negative trail length gives higher stability at lower speeds. Though being successful at lower speeds, at which the manual control is taken over by autonomous control, this concept fails at higher speeds because at higher speed the control still has to be manual and the bike remains no special than normal bikes. Such a concept cannot be taken further to make the bike autonomous as it is obsolete when the bike needs to turn at higher speed and dodge an obstacle.

This brings back to the traditional theories to self-balance the bike using the gyroscopic force through the action of the flywheel. The value of gyroscopic force is given by the cross product of the angular momentum of body and angular velocity of the axis of rotation. This concept is applied in a bike by introducing a rotating flywheel inside the bike and changing its axis of rotation in accordance with the desired force. Today, the bikes cannot be balanced unless the biker controls the handle. This led the authors to the making of a self-balancing bike.

A common misconception about the working of the bike that people has is that only handle is rotated to take a turn. This does not explain how a person riding a bike without holding handle is able to turn. Also, it has been experimentally proven that unless the bike is tilted, simply turning the handle would result in toppling of vehicle due to outward force, and this is the reason why race bikers have to tilt by a great amount to balance while turning. The explanation of turning of the biker without holding the handle of the bike can be given by thinking the process other way round. The bike actually is tilted first. This produces a gyroscopic force on the wheel causing it to turn, and thus restoring the bike to the mean process. This observation is used in the project to make the self-sufficient bike turn. A precisely tilted bike would result in a precise turn for a given speed. Such a precise tilt will be achieved by the principle of conservation of angular momentum coupled with gyroscopic phenomena to maintain the bike stability simultaneously.

Self-balancing bike could balance using gyroscopic force and can be further extrapolated to detect the location and path through GPS to reach the desired location, thus making it self-sufficient. So, in this paper, authors present the two-wheeled bike which can balance itself with the help of gyroscope and can make a safer transportation. This will reduce accidents and injury with better control on uneven and hilly terrains while driving two-wheelers. It also proves to be helpful for differently abled people. Such a bike solves the problems of uncontrolled traffic, accidents, and offers a leisure full and safe ride. The bike, being small and very much maneuverable, has an edge over other forms of transport. It would save human time, efforts, and set a stress-free situation.

2 Mechanical Design and Construction

Vehicles like ships and aircrafts also use the gyroscopic effect as an external disturbing couple is always acting on such vehicles. To neutralize the disturbing couple, an equal and opposite reactive couple is applied, which is necessary for its stability.

The rate of change of angular momentum of the flywheel induces a reactive torque in response to that caused when the bike is imbalanced. When the spinning axis rotates itself, it causes the prototype to rotate in the direction opposite to the fall direction, thus neutralizing the disturbance and stabilizes the prototype. During the design process, a key factor was to make such a flywheel that can induce enough torque with the constraints of motor speed as well as servo speed.

This paper proves that a body supported in a single axis under a 3-D atmosphere (like an inverted pendulum) can be balanced using the action of a gyroscope.

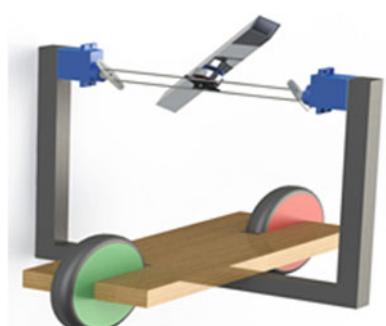
The concept of active control system inducing gyroscopic force is used. The model is made robust and an aluminum rod is used as a flywheel to increase inertia. Holes were made in the end to add nuts and bolt in order to increase the inertia of the wheel as per requirement. The radio control system is removed and the bike is controlled by onboard microcontroller with embedded gyro sensor [2]. Figure 1 shows the prototype model using radio control system.

Following are the parts used to create the prototype (Fig. 2):

Fig. 1 Final model



Fig. 2 Prototype
Solidworks model



1. Carbon fiber frame	- height: 15 cm width: 18.5 cm
2. Wooden base	- length: 30 cm width: 7 cm
3. 12 V DC brushless motor	- 750 kV and 3 phase
4. Electronic speed control	- 20 A
5. Aluminum bar	- 15 cm
6. Flathead stud	
7. Plastic wheels	- radius: 3.5 cm
8. Arduino UNO	
9. Servomotors	- speed: 0.12 s/60°
10. MPU 6050	- accelerometer and gyroscope

3 Control System

Control system in bike consists of following hardware components (Fig. 3):

- (1) **Arduino Uno:** For controlling a dynamic system accurately, it is important to acquire data at high rate. Arduino can support 115,200 bits per second. Currently, the arduino is acquiring data from MPU6050 to calculate the angle by which the bike has been tilted from mean position. This angle is termed as an error signal and is fed into PID controller. The controller then determines the speed and direction of rotation of servomotor, thus generating torque. The torque is generated accurately to balance the bike.

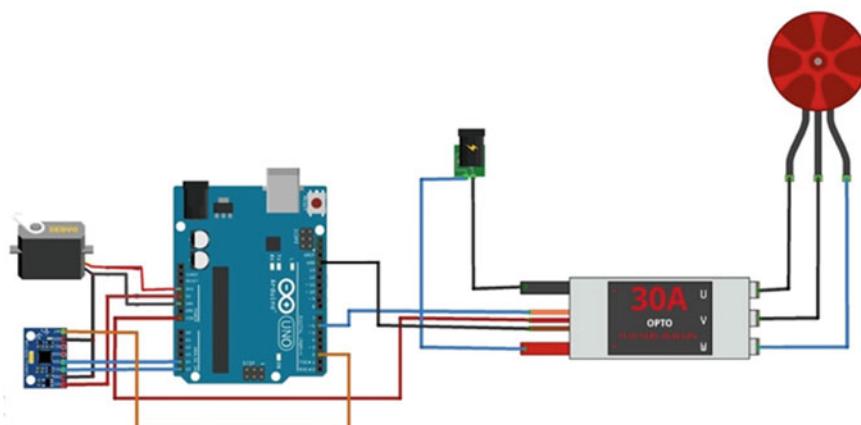


Fig. 3 Final design circuit diagram drawn in *Fritzing*

- (2) **MPU6050:** MPU6050 consists of three-degree accelerometer and three-degree gyroscope sensor which result in accurate tilt angle determination. MPU 6050 works as feedback unit in closed-loop system (Fig. 5), taking tilt angle as input and giving the feedback to Arduino.
- (3) **DC Brushless motor:** It is Rated at 12 V, 750 kV. The flywheel of the DC brushless motor rotates at predefined rpm.
- (4) **Servomotors:** These motors have 19.6 kg/m of stall torque and operate at 6 V.

A block diagram for the bike control system is shown in Fig. 4 depicting the control units and signal flow in a closed loop. Speed of servo is controlled as soon as an error is observed by gyro sensor as it deviates from set point.

Signal flow diagram of bike, as shown in Fig. 5, starts with microcontroller taking data from mpu6050 and making decisions based on the statement defining upright position. To achieve this flywheel is rotated through calculated amount and to given position. This will generate gyroscopic force which tries to oppose the action which induces it. The loop works till the bike attains upright position.

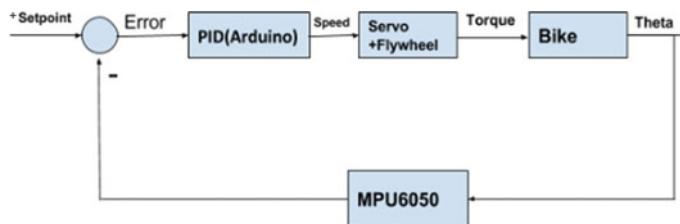


Fig. 4 Block diagram of bike

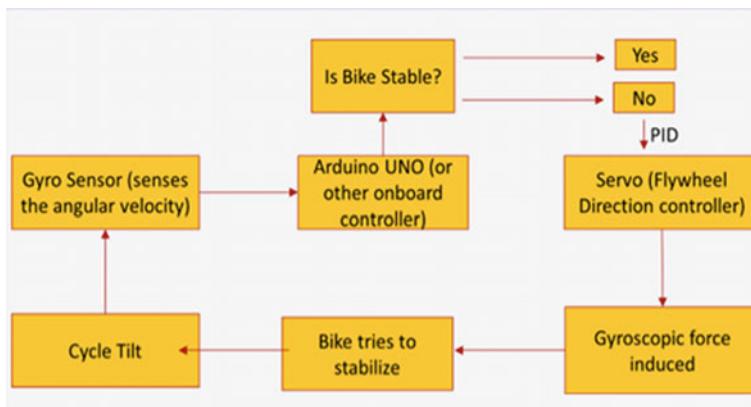


Fig. 5 Control flow chart

4 Experimental Results

Experiment is conducted and the qualitative dependence of induced torque with respect to angle of flywheel as well as angular velocity of servo is determined. Figure 6 shows the variation of torque component responsible to balance the prototype with angle of flywheel. The mathematical relation that the graph follows is given as follows:

$$\tau = (I * \omega_1) * \omega_2 * \sin \theta * \sin \beta$$

where

τ = Induced torque,

I = Moment of inertia of flywheel,

ω_1 = Angular velocity of flywheel,

ω_2 = Angular velocity of servo,

Θ = Angle between ω_1 and ω_2 which is $\frac{\pi}{2}$,

β = Angle of flywheel with the horizontal plane.

In the prototype, Θ has been fixed to 90° to induce maximum torque and the variation of torque with β is observed. As servo rotates, angle of the net torque with the horizontal plane varies. This angle is equal to the angle swept by the servo, due to which useful torque varies. This variation is nonlinear more precisely sinusoidal.

Figure 7 shows the variation of normalized torque with angular velocity of servo (ω_2), keeping other variables constant and setting β at 90° . This means that the torque observed is plotted when the flywheel passes through the mean position, i.e., when β is 90° .

Fig. 6 β (radians) versus normalized torque

TORQUE COMPONENT IN THE DIRECTION OF CYCLE MOTION

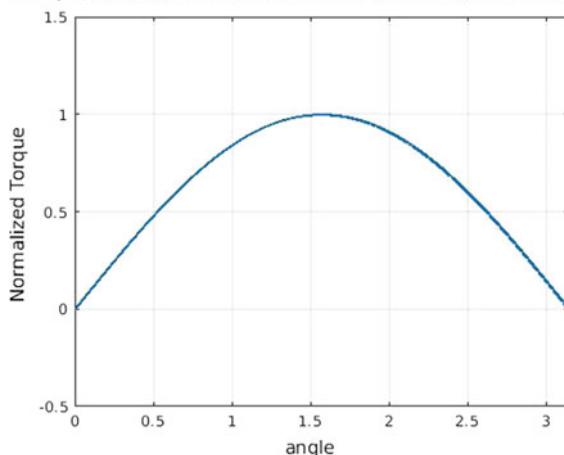
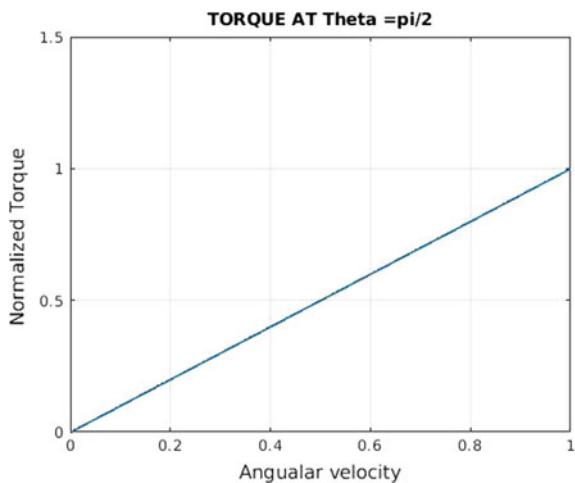


Fig. 7 ω_1 versus normalized torque



In Ziegler–Nichols method, first all PID constants are set to zero and then are increased till the bike starts to have infinite constant oscillations in time as given in Fig. 8 [3]. Following procedure has been conducted to deduce the values of K_p , K_i , and K_d :

STEP 1 The value of K_p is changed using trial and error method till constant oscillations are achieved.

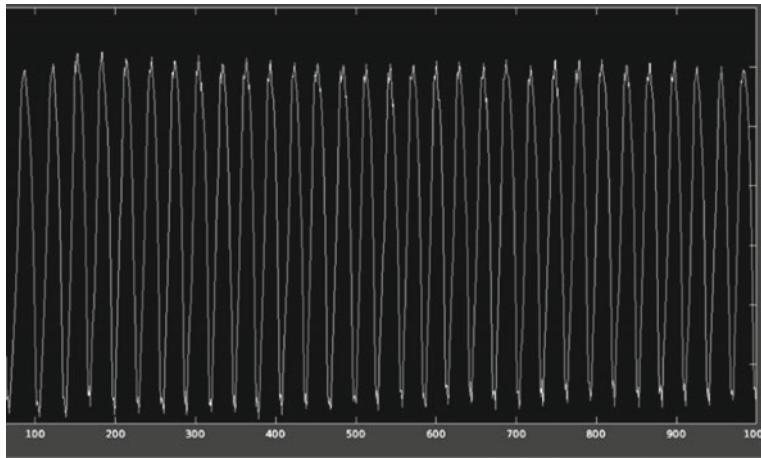


Fig. 8 Angle oscillations in time of bike

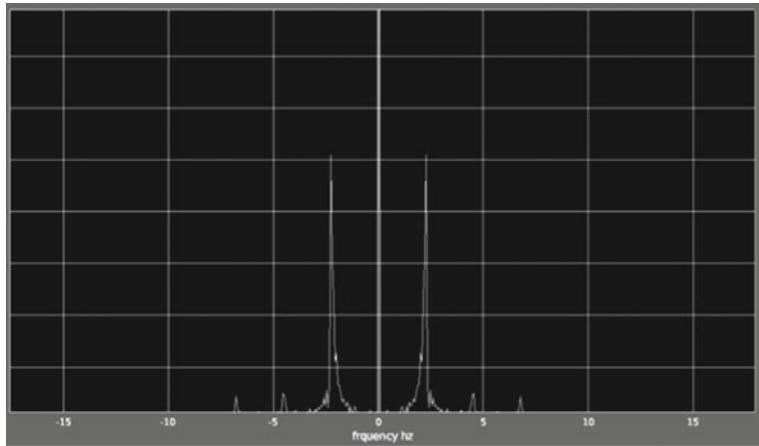


Fig. 9 Fourier transform of angle variation

- STEP 2** Figure 8 depicts the oscillations with respect to time. To get the time period of these oscillations, a Fourier transform is taken. Fourier transformation gives the relative amount of every frequency detected and shows in the form of amplitude of frequency versus the value of frequency.
- STEP 3** Figure 9 shows the Fourier transformation of the same. A peak value of frequency is obtained (the other peak is the same frequency in other direction) from which the time period is calculated.
- STEP 4** The calculated time period is used to calculate K_i and K_d using following formula:

$$K_p = 0.6K_u$$

$$K_i = \frac{2 * K_p}{T_u}$$

$$K_d = \frac{K_p * T_u}{8}$$

From plots following values are observed:

$$T_u = 0.5\text{s},$$

$$K_p = 63$$

Using above values and equations given below, the PID constants are determined to balance the bike.

$$K_p = 37.8$$

$$K_i = 151.2$$

$$K_d = 2.33$$

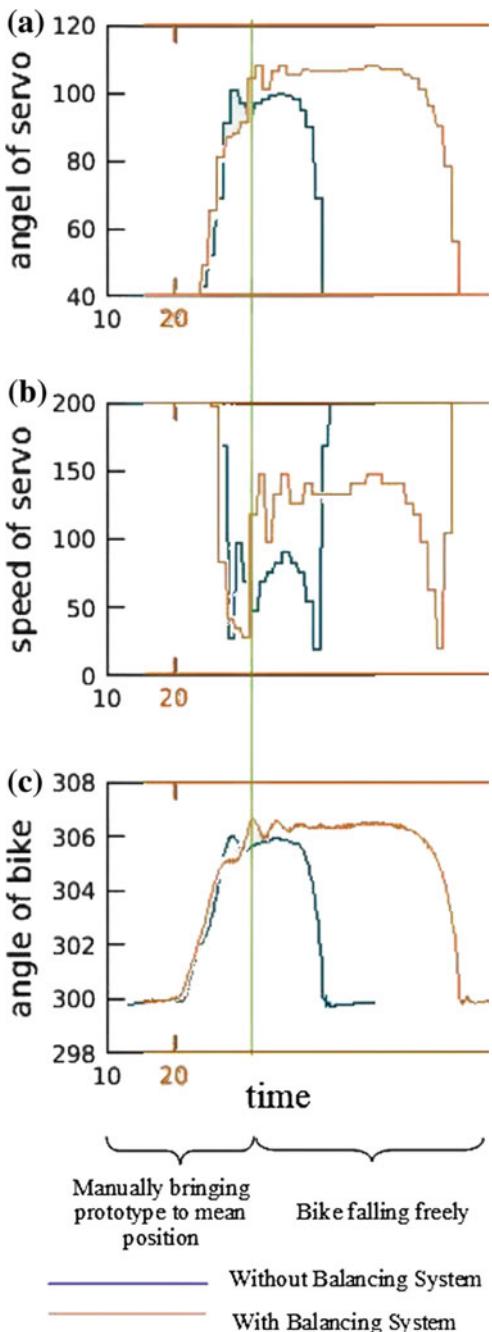
In the algorithm used, as shown in Fig. 10, deviation in angle of bike from upright position (306.2° as shown in graphs) results in rotation of servomotor by the same proportion to angle. As angle of servo changes (reference being the 90°), its angular velocity increases so as to resist the change. The last graph shows how fall of bike is different in two cases, one in which our controller works and other in which it is switched off. It is clearly visible that with controller, the bike is balanced for a longer duration [4].

5 Conclusions

This paper presents a design of the prototype of a two-wheeler self-balancing vehicle. The prototype remains balanced on the two wheels without any support from the external agent [5]. A mild external force applied to disbalance the vehicle can also be overcome. The system proposed boasts of having the capability of reducing accidents, controlling traffic, and driving in difficult terrains making it a much safer ride.

A many more safety features are incorporated in the bike making it more reliable, such as obstacle avoidance, path detection, and Bluetooth, and especially the gyro-stability system which keeps the vehicle upright even during any collision. It is as compact as normal bike, requiring very less space for parking. Cost-effectiveness factor is another criteria considered while designing the vehicle. As a result, it is a power efficient, indigenous, and economic transportation system as compared to other vehicles, serving as a panacea for the two-wheeler transportation.

Fig. 10 **a** Rotational angle of servomotor as bike begins to fall, **b** Angular velocity of the bike as bike begins to fall, **c** Change in angle of bike as it falls (306.2 being the upright position)



References

1. Govardhan, P., et. al.: Survey on self balancing two wheel electric prototype. Int. J. Eng. Res. General Sci. **5**(5) (2017). <http://oaji.net/articles/2017/786-1509075662.pdf>
2. https://drive.google.com/open?id=1_bVxmLYIWAcfpckaNb6W3s0mCwfVJJ0B
3. <https://drive.google.com/file/d/1RSPIBUfq5gyxGHgj8hbALNmvpM-P8GI2/view?usp=sharing>
4. <https://drive.google.com/file/d/1sM2Lf-a-RPIzHs1Yyh1kkyupbGDjsvrOn/view?usp=sharing>
5. <https://drive.google.com/open?id=1BdMi7oawPHJMsNUchn9hRUM-nEPN-obt>

Parallel DES with Modified Mode of Operation



Kinjal Chaudhari and Payal Prajapati

Abstract In this paper, we tried to model parallel version of Data Encryption Standard (DES) by means of the modified mode of operation. Block size specifications followed by DES have been adopted too. Design configuration has shown serialized intermediate key generation and parallelized operations on plaintext blocks, which indicate faster encryption as compared to the standard one. We have performed various statistical tests suggested in the literature for conforming security of the proposed model. Also, case studies on potential vulnerabilities have been shown. Comparative analysis with the standard DES has revealed that the proposed parallel DES is more secure due to inter-block data dependency of keys and serialized decryption besides faster encryption. We have concluded our work with future directives.

Keywords DES · CBC · Cryptography · Parallel encryption

1 Introduction

The degree of confrontation to, or fortification from, damage is known as security. It can be applied to any vulnerable or valuable asset. Securing information is essential with the increased growth of high-profile attacks using advanced technologies. In the field of security, data encryption standard, accepted as DES [34], has widespread usage [6, 31]. However, the potential vulnerability of DES to a brute-force attack enforced development of approaches with multiple DES encryptions and multiple keys, which are known as double DES (2DES) and triple DES (3DES or TDES) [32]. These are block ciphers and they can be further applied in a variety of applications

K. Chaudhari · P. Prajapati (✉)

Institute of Technology, Nirma University, Ahmedabad 382 481, Gujarat, India
e-mail: payal.prajapati@nirmauni.ac.in

K. Chaudhari
e-mail: kinjal20c@gmail.com

using modes of operation defined by National Institute of Standards and Technology (NIST) [26].

DES using standard 64-bit plaintext blocks and a 56-bit key undergoes 16 rounds of repetition; here, output of one round, i.e., a 64-bit block becomes input to the next round. When DES is applied to Cipher Block Chaining (CBC) mode, output of the 16th round, i.e., a 64-bit ciphertext block, is given as an input for operating the next 64-bit plaintext block along with key. This ensures that even if a plaintext block is repeated, encryption with the same key would provide different ciphertext blocks. Hence, it introduces serialized behavior where output of one block is given as an input to the next block when applied to CBC. We have modified this mode of operation to achieve parallelization for DES. Our aim is to achieve faster encryption with comparable security as that of the standard DES. We firmly believe that the proposed serialized decryption is the main reason to achieve comparable security as it introduces complexity to cryptanalysts.

1.1 Major Contributions

We can summarize our contributions as follows:

- We have modified CBC mode of operation to make it work in parallel.
- We have applied the standard DES encryption in this modification and discussed the parallel approach.
- We have also demonstrated serialized decryption of DES in our work.
- We have compared our approach with the existing ones in terms of security and encryption speed.
- We have analyzed the robustness of our approach against various attacks.

The remaining paper is broadly organized as follows: Sect. 2 provides a brief introduction to the existing approaches. Section 3 includes our proposed work where we discuss parallelization in encryption and serialization in decryption of standard DES using modified CBC. We have statistically analyzed our model and discussed case studies for its robustness against various vulnerabilities in Sect. 4. We have concluded our work with future scope in Sect. 5.

2 Related Work and Concepts Used

In this paper, we have targeted DES algorithm with one of the modes of operation (CBC) where it can be utilized. We have modeled DES with modified CBC so as to have parallel encryption.

Our literature survey includes some of the recent literature where authors introduced parallel version of symmetric key encryption algorithms. All these methods were based on identifying the most time-consuming elements from the source code

of the cryptographic algorithms, for example, loops. The well-known techniques for making loops parallel were used to speed-up the execution of loops which were directly responsible for encryption and decryption processes, thereby achieving faster encryption. They have majorly used the OpenMP Application Program Interface (API) for parallel implementation.

Authors in [2] presented parallel DES using the OpenMP standard. They found most of the “for” loops of DES algorithm to be well suited for parallel implementation. They evaluated speed-up of their experiment as compared to the sequential standard DES. Authors of [3] incorporated parallel DES suggested in [2] with various modes of operation. They also parallelized source code of mode of operation by separating independent and dependent implementation parts. However, the inherent sequential nature of modes of operation remained there, making all the plaintext blocks being encrypted in serial fashion. Authors of [12] felt that even though good speed-up has been achieved in response to all the efforts to parallelize the existing conventional cryptographic algorithms with hardware techniques [11, 14, 24, 25, 37] by building special-purpose microprocessors or with software techniques, they cannot be fully parallelized or implemented efficiently due to the dependency problems and the inability to efficiently modularize the sections of the algorithms hover around and haunt the parallelization. So they developed a different class of cryptographic algorithm involving modular arithmetic concepts which have achieved good speed-up along with better security. On the other hand, authors in [17] implemented Advanced Encryption Standard (AES) [8] using CUDA platform. They implemented a parallel encryption algorithm using GPU parallel computing and its optimized design for cryptography.

We majorly found papers on software- and/or hardware-based parallel implementations of part of the DES algorithm. In this paper, instead of sequential encryption using DES with CBC, we have proposed parallel encryption of the plaintext blocks using DES when applied with CBC; this is performed by modifying the inherent chaining concepts of this mode of operation. This modified CBC can be used with any symmetric key cryptographic algorithm. As DES has been widely used in practice, we have applied this modified CBC with DES to have simultaneous encryption of plaintext blocks. As per our survey, we did not find research oriented to modifying mode of operation for receiving parallel DES implementation. In this section, we discuss the concepts used in our model.

2.1 Concepts Used

In our proposed approach, we have used DES and CBC mode, which are briefly explained as follows.

Data Encryption Standard. It is a symmetric key encryption algorithm [34]. Here, data of 64-bit plaintext blocks are encrypted using a 64-bit key to generate 64-bit output blocks. Single plaintext block in the standard DES proceeds in three phases:

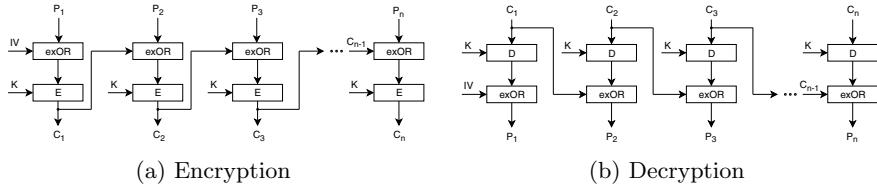


Fig. 1 DES in CBC

initial permutation (IP), 16 rounds of the same function, and inverse initial permutation (IP^{-1}). A 64-bit input key is passed through a permutation function ($PC-1$) that generates a 56-bit output. For each of the 16 rounds, a different subkey is produced from it by repeated left circular shifts and permutation ($PC-2$). If the data to be encrypted is divided into more than one plaintext block, these same 16 subkeys are reused for encrypting each of the next plaintext block(s). Decryption takes place in a similar manner in DES.

To resist against the potential vulnerabilities of DES [4, 6, 7, 19], new alternatives were found. A completely new algorithm, named AES [8], was introduced as an alternative to DES; however, multiple times encryption using the same DES was proposed as the other alternative way. Security of multiple encryptions has been discussed in [22]. The concepts of DES can be understood using its simplified version, viz., S-DES as given in [28].

Modes of Operation. A mode of operation is a technique for enhancing the effect of a cryptographic algorithm or adapting the algorithm for an application. Four of the basic modes are Electronic Codebook (ECB), Cipher Block Chaining (CBC), Cipher Feedback (CFB), and Output Feedback (OFB); Counter (CTR) mode was included then after [10]. Any encryption technique can be applied to any of these modes [26]. In Fig. 1a, b, we have demonstrated DES encryption (E) and decryption (D), respectively, using one of the modes, viz., CBC; here, the i th plaintext block gets preceding ciphertext block for the EX-OR operation. Similarly, DES can be given as the encryption algorithm in any other mode.

Here, because of the block chaining, the serialized behavior of DES continues when applied in any mode of operation except for ECB and CTR. It does not give faster outcomes, instead, the dependency on the preceding ciphertext block slows down the speed. In our proposed work, we have tried to break the serialized behavior by breaking the chain between output of one block and input to the next block. The same has been discussed in further sections.

3 Proposed Approach

We propose parallelization of DES with modification in the mode of operation. We have considered the input block size of 64 bits, i.e., the plaintext has been divided into 64-bit blocks (P_1, P_2, \dots, P_n) to generate 64-bit ciphertext blocks (C_1, C_2, \dots, C_n).

Here, IV stands for initialization vector, which we assume being securely shared between both, sender and receiver; other specifications may be desirable for selection of IV [36]. We have indicated encryption and decryption using 16 rounds of DES by E and D , respectively. The concept of block chaining has been utilized in our proposed work.

3.1 Parallel Encryption

We provide a 64-bit secret key, K and 64-bit IV as inputs to an EX-OR function to generate K_1 as shown in Fig. 2. K_1 is then EX-ORED with plaintext block, P_1 to generate an intermediate value, Y_1 . Here, Y_1 serves as input to the next block. Ciphertext block, C_1 , is the outcome of DES algorithm applied on P_1 and K_1 . For any i th block, encryption takes place as given by Eqs. (1)–(3).

$$K_i = K \oplus Y_{i-1} \quad (1)$$

$$Y_i = P_i \oplus K_i \quad (2)$$

$$C_i = E(K_i, Y_i) \quad (3)$$

Here, instead of passing C_{i-1} to the next, i th block like that is shown in Fig. 1a, we have passed an intermediate value, Y_{i-1} to the i th block. Also, we have not passed any Y_{i-1} value directly to the P_i block and we have EX-ORED it with the secret key K to generate K_i , which is then EX-ORED with P_i and also given as input key to DES encryption, E . This feature yields out that for every parallel encryption, and similarly for decryption, we have new key(s) with which DES would generate 16 different subkeys for each plaintext block, and similarly for each ciphertext block. Hence, we do not have the same 16 subkeys in any case.

The key generation takes place in a serialized way whereas the remaining encryption algorithm(s) for each plaintext block has been parallelized. In Fig. 2a, the horizontal dashed line differentiates between the serialized and parallelized portions of DES encryption, i.e., the upper half and the lower half, respectively, with a modified mode of operation.

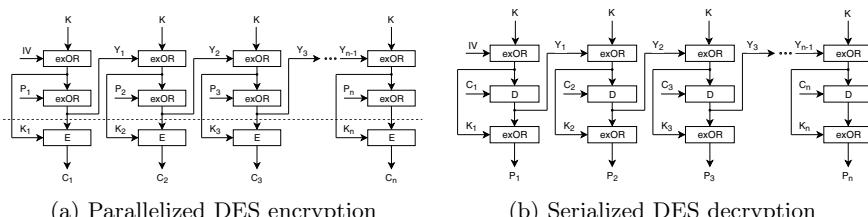


Fig. 2 Parallel DES encryption and decryption

3.2 Serialized Decryption

A cryptographic algorithm should have simplicity in design; however, more complexity in cryptanalysis is also an important aspect that can strengthen it against attacks. Here, we discuss the decryption part of our proposed work, which provides serialization instead of parallelization like that in encryption. Figure 2b depicts the decryption of our proposed work for DES. For any i th block, decryption can be given by Eq. (4).

$$P_i = K_i \oplus D(K_i, C_i) \quad (4)$$

4 Analysis

Our approach for having parallel encryption using DES must be statistically tested. The analysis of such tests and case studies on possible attacks are explained in this section. As per our survey, there are no specific tools available for checking the security of a cryptographic algorithm. However, we have come across several papers that have used different measures to evaluate the security and performance of their proposed approach [9, 13, 16, 23].

4.1 Statistical Tests

Security of an algorithm is dependent on various parameters such as block size, key size, diffusion, and confusion [1]. We have considered two properties, namely, diffusion and confusion as suggested by Claude Shannon [29, 30] for statistically testing parallel DES. To ensure that parallel DES is cryptographically secure, we have performed the following tests.

Diffusion. It is the method in which the statistical structure of the plaintext is dissipated into long-range statistics of the ciphertext [32]. Here, the redundancy of the plaintext is increased such that each plaintext letter (digit) affects many ciphertext letters (digits); the statistical relationship between the plaintext and corresponding ciphertext must be as complex as possible.

As given in [5, 20, 21, 27, 33], if a ciphertext is having a good diffusion property then flipping one bit of the input changes every bit of the output with a probability close to $\frac{1}{2}$. This is known as the avalanche effect [32]. It is a desirable property of any cryptographic algorithm. Hence, change in one bit in either the plaintext (strict plaintext avalanche criterion (SPAC)) or the key (strict key avalanche criterion (SKAC)) should produce a significant change in approximately half of the bits in the ciphertext. Analysis becomes more difficult while mounting an attack on such algorithm [1].

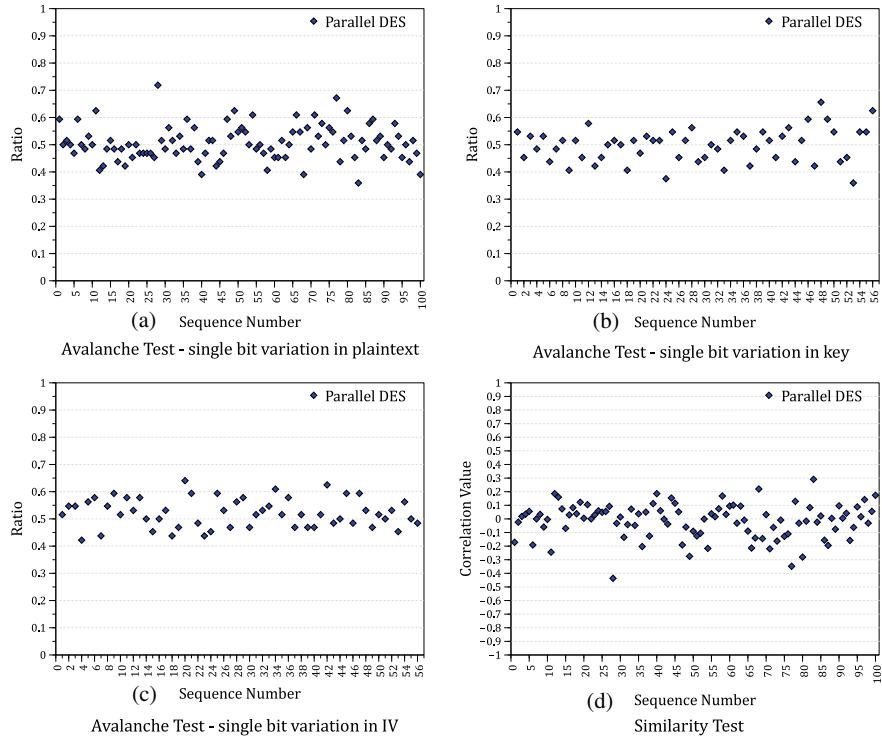


Fig. 3 Statistical tests: **a, b, c** Avalanche tests; **d** Similarity test

In our approach, we have two input seed values, IV and secret key, K . We have tested the avalanche effect with three variations in plaintext, key, and IV . We have used 100 64-bit samples of plaintexts and evaluated their corresponding ciphertexts with fixed inputs, key, and IV . We randomly flipped one bit in each of the plaintexts and derived their corresponding ciphertexts with fixed inputs as earlier. The avalanche effect provides a ratio of flipped ciphertext bits to the total. Here, Fig. 3a indicates sample plaintext sequences and respective ratio when only one bit is flipped randomly in each plaintext. Similarly, for the fixed plaintext, we have used 56 variations of a key by flipping only one bit in each execution; 56 variations of IV by flipping single bit in each encryption are performed as well. Figure 3b, c show avalanche effect of ciphertexts generated using 56 variations of a key and 56 variations of an IV , respectively. In our experiment, the average avalanche effects achieved for plaintext, key, and IV variations are 0.5066, 0.4975, and 0.5218, respectively. Hence, it indicates that parallel DES satisfies the diffusion property.

Confusion. It makes the statistical relationship between the plaintext and the ciphertext as complex as possible. This must be performed so as to thwart attempts to deduce the encryption key used [32].

The dependency of the individual output bits on the input bits can be given by the correlation coefficient. The correlation values can determine the confusion effect of the block cipher; they lie between -1 and 1 and measure the degree of linear relationship. Independent variables give correlation equal to 0 . The correlation values are considered to have weak positive or negative linear relationship if they are between 0 and 0.3 or -0.3 and 0 , respectively [1].

As explained earlier, we have taken 100 plaintext samples and evaluated their corresponding ciphertexts for the fixed inputs. We have applied the similarity test on these sequences to find correlation. As shown in Fig. 3d, we achieved the correlation values approximately between -0.3 and 0.3 which indicates weak linear relationship. Hence, our approach provides a significant level of confusion.

4.2 Case Study on Attacks

Here, we discuss case studies on vulnerabilities of parallel DES and evaluate bit security levels. The bit security estimates the number of computational steps or operations required to reveal secret key, for example. In our case study, the seed values, IV , and the secret key, K are assumed to be shared between the sender and the receiver in a secure manner prior to communication and are unknown to the attackers.

Brute-Force Attack. The cryptanalyst tries out every possible key until an intelligible translation of the ciphertext into plaintext is obtained. On average, half of all possible keys must be tried in a brute-force attack to succeed [32].

In brute-force attack, a cryptanalyst has access to the ciphertext. For the standard DES of 64 -bit plaintext and 56 -bit secret key, this attack can be performed using 2^{56} combinations to find the secret key [32]. As shown in Fig. 1a, each encryption uses the same secret key for each plaintext block. This states that for an attacker, a maximum of 2^{56} combinations would provide the secret key used and the same key can be applied to other ciphertext blocks in that sequence.

As shown in Fig. 2a, an intermediate key, K_i , is given to DES encryption of the i th plaintext block. The secret key, K , is used in each block for generating the intermediate values such as K_i and Y_i ; however, it is not given directly to DES. For 64 -bit data blocks in parallel DES, brute-force may be applied on a 64 -bit intermediate key; it seems to require trying out 2^{64} possible combinations to deduce the key used. However, every block is encrypted using a different intermediate key, and therefore the deduced key cannot be used with other ciphertexts to derive plaintext. Thus, performing a brute-force attack on intermediate keys would require trying out 2^{64} combinations for each block, which is practically not advisable. Whereas in the standard DES with CBC, the previous ciphertext, C_{i-1} , for example, is given as an IV to the next, i.e., i th round. The attacker has the ciphertext value available. Here, revealing only K through single DES block would be sufficient. Thus, a maximum of 2^{56} combinations should be tried.

Another way of performing this attack on parallel DES is to reveal the secret key, K and IV values, based on which other intermediate keys can be generated. Assuming each of these seed values to be 64-bit long, a maximum of 2^{128} combinations need to be tried out. The importance of not using seed values interchangeably is due to the next blocks where secret key, K , is also given as an input, and interchanging the seed values would not lead to a successful attack. In other words, every intermediate key, K_i , is generated using plaintext, P_i along with K and Y_{i-1} . To reveal the secret key, K , we must try every possible Y_i , making the number of combinations as large as 2^{64} for one K , and hence, a total of 2^{128} possible combinations. This shows that the bit security of brute-force attack on parallel DES, 128 bits, is significantly larger than that of the standard DES.

Known Plaintext Attack. Here, along with known encryption algorithm and ciphertext, one or more plaintext–ciphertext pairs are also known to the cryptanalyst that were formed with the secret key [32, 35]. Literature [18] has shown that properties of the standard DES can be exploited using the known plaintext attack and it is considered as the best-known attack. So far, the best-known attack on DES requires 2^{43} known plaintexts which was performed by Matsui's linear cryptanalysis [18]. Also, an attack was reported which requires 2^{47} chosen plaintexts through differential cryptanalysis [15].

Though we have not performed linear and differential cryptanalysis on parallel DES, we tried meet-in-the-middle attack using known plaintext–ciphertext pairs. As can be seen from Fig. 2a, any intermediate key, K_i , used to decrypt ciphertext block, C_i , and the same key in EX-OR operation with Y_i to generate plaintext block, P_i , i.e., $D(C_i, K_i)$ and $K_i \oplus P_i$ provide the same value, Y_i . This fact allows to perform meet-in-the-middle attack which is performed as follows: an attacker has tried to find out a successful value of Y_i which means that there is at least one corresponding K_i available such that $K_i \oplus P_i$ and $D(C_i, K_i)$ would provide the chosen Y_i for any one of the plaintext–ciphertext pairs available. The probability of the chosen Y_i to be successful is $\frac{m}{2^{64}}$, where m is the number of known (P, C) pairs. A basic result from probability theory is that if the balls are not replaced, the expected number of draws required to draw one red ball out of a bin containing m red balls and $N - m$ green balls is $\frac{(N+1)}{(m+1)} \approx \frac{N}{m}$, where value of N is 2^{64} [32].

To break our approach, the observation is that an attacker requires at least two consecutive known (P, C) pairs along with the internal order of their occurrence. To reveal K , one must have the value of Y_i and K_{i+1} . Assuming the scenario where we have two non-consecutive (P, C) pairs, for example, (P_2, C_2) and (P_4, C_4) . As we discussed earlier, we can go with finding Y_2 and Y_4 using the (P_2, C_2) and (P_4, C_4) pairs, respectively, with an effort of 2^{64} each, giving total of 2^{65} operations. This will give the corresponding K_2 and K_4 . However, to derive the secret key K , an attacker would require K_3 and Y_2 or K_4 and Y_3 . Due to the unavailability of K_3 and Y_3 , an attacker must assume 2^{64} possible combinations for any of these, which would give 2^{64} possible K , which is of no use. If the attacker could have any of K_3 or Y_3 , he/she could derive single value of K , which is only possible in the case of consecutive

known (P, C) pairs. To assess validity of the deduced K , one can try it against other known (P, C) pairs.

Besides having knowledge of two consecutive pairs, the attacker also requires internal order of the plaintext sequences. This observation is explained by a scenario where we have two (P, C) pairs available, namely, (P_2, C_2) and (P_3, C_3) ; however, we do not know the internal order. Suppose, an attacker tries (P_2, C_2) as the first pair and (P_3, C_3) as the second pair, which is in fact the correct order, he/she would try to find out Y_2 , Y_3 and their respective K_2 and K_3 as explained above with a total effort of 2^{65} . But if an incorrect order is selected, then even after performing 2^{65} operations, the attacker would not get the correct secret key, K . Hence, in the absence of knowledge of the internal order, the attacker has to try all possible order combinations as well, i.e., two in this scenario. Complexity would be much greater when a higher number of known (P, C) pairs are available.

These can be explained by a general scenario where the attacker is assumed to have knowledge of m (P, C) pairs available; however, the internal order and consecutiveness of the pairs are unknown. Here, the attacker would have to try out every (P, C) pair with every other pair, maintaining the order to see if any of the deduced key, K , works well for other known pairs. This may lead to a total of $m \cdot (m - 1)$ trials, each of which requires to have 2^{65} operations. If none of the available pairs is consecutive, the attacker would not get the valid K even after these operations. The probability of having a successful attack would be higher in case of having larger m as the probability of getting at least one consecutive pair would be high, of course at the cost of more number of operations.

Exploiting Repetition of Plaintext Blocks. One of the main advantages of CBC is that for repeated plaintext blocks, the ciphertext would be different due to block chaining. We have analyzed our approach to exploit repetition of plaintext blocks. It has been observed that in the case of three or more consecutive same plaintext blocks, the respective ciphertext blocks get repeated in an alternative manner. For example, for three consecutive same plaintext blocks, the first and the third ciphertext blocks would be the same. Similarly, for four consecutive same plaintext blocks, the first and the second ciphertext blocks would be same as the third and the fourth, respectively. One may exploit this repetition of ciphertexts to reveal plaintext or key. However, this would also require to try 2^{65} to deduce the key.

5 Concluding Remarks and Future Scope

In our work, we have studied modes of operation and have modified CBC to perform parallel DES encryption; decryption, on the other hand, is serialized in our approach. The parallelized encryption increases the speed-up, whereas serialized decryption increases complexity for the cryptanalyst to break the approach.

We have shown the statistical test of our proposed model. Two of the most desirable properties of any cryptographic algorithm, diffusion and confusion, are achieved by

parallel DES. These properties are tested using avalanche effect and correlation metric and the results have determined that parallel DES is capable of frustrating the cryptanalysts by providing complex relationships of plaintexts with ciphertexts and keys. We may note that satisfying diffusion and confusion properties has proven comparable security of parallel DES.

The standard DES of 64-bit blocks can be exploited using the brute-force attack with 2^{56} combinations. In parallel DES, the brute-force attack requires 2^{128} combinations even with 64-bit blocks. We have designed it to be scalable; the block size can be increased in parallel DES. Considering 128-bit blocks, our approach is capable of making the brute-force attack as large as 2^{256} combinations. We have also explained known plaintext attack with various possible scenarios. The observations indicate that knowledge of consecutive (P, C) pairs along with their internal order are requirements for having a successful known plaintext attack; however, for the block size of 64 bits, a cryptanalyst is expected to try a total of 2^{65} combinations to deduce the secret key. These values for security levels are significantly larger than that of the standard DES.

Though S-boxes of the standard DES have been analyzed on the basis of linear and differential cryptanalysis, the same may be applied in our approach as we have not modified DES in particular. We have not shown linear and differential cryptanalysis on parallel DES; however, interested researchers may try to exploit its properties by trying various attacks on our approach. One may also wish to implement parallel DES on APIs such as OpenMP to make it work in the more efficient way and provide speed-up.

References

1. Alabaichi, A., Ahmad, F., Mahmood, R.: Security analysis of blowfish algorithm. In: 2013 Second International Conference on Informatics and Applications (ICIA), pp. 12–18. IEEE (2013)
2. Beletskyy, V., Burak, D.: Parallelization of the data encryption standard (DES) algorithm. In: Enhanced Methods in Computer Security, Biometric and Artificial Intelligence Systems, pp. 23–33 (2005)
3. Bielecki, W., Burak, D.: Parallelization of standard modes of operation for symmetric key block ciphers. In: Biometrics, Computer Security Systems and Artificial Intelligence Applications, pp. 101–110. Springer (2006)
4. Biham, E., Shamir, A.: Differential Cryptanalysis of the Data Encryption Standard. Springer Science & Business Media (2012)
5. Chandrasekaran, J., Subramanyan, B., Raman, G.: Ensemble of blowfish with chaos based s box design for text and image encryption. *Int. J. Netw. Secur. Its Appl.* **3**(4), 165–173 (2011)
6. Coppersmith, D.: The data encryption standard (DES) and its strength against attacks. *IBM J. Res. Dev.* **38**(3), 243–250 (1994)
7. Courtois, N.T., Bard, G.V.: Algebraic cryptanalysis of the data encryption standard. In: Proceedings of the 11th IMA International Conference on Cryptography and Coding, pp. 152–169. Springer-Verlag (2007)
8. Daemen, J., Rijmen, V.: The design of Rijndael: AES-the Advanced Encryption Standard. Springer Science & Business Media (2013)

9. Doganaksoy, A., Ege, B., Koçak, O., Sulak, F.: Cryptographic randomness testing of block ciphers and hash functions. *IACR Cryptol. ePrint Arch.* **2010**, 564 (2010)
10. Dworkin, M.J.: Recommendation for block cipher modes of operation: Galois/counter mode (GCM) and GMAC. Technical report (2007)
11. Ichikawa, T., Kasuya, T., Matsui, M.: Hardware evaluation of the aes finalists. *AES Candidate Conf.* **2000**, 279–285 (2000)
12. Jose, J.J.R., Raj, D.E.G.D.P.: PACMA—An adaptive symmetric cryptographic algorithm for parallel computing environments. In: *The Proceedings of the Fifth International Conference on Advances in Recent Technologies in Communication and Computing*, Bangalore, India (2013)
13. Juremi, J., Mahmud, R., Sulaiman, S.: A proposal for improving AES S-box with rotation and key-dependent. In: *2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic (CyberSec)*, pp. 38–42. IEEE (2012)
14. Kim, H.W., Lee, S.: Design and implementation of a private and public key crypto processor and its application to a security system. *IEEE Trans. Consum. Electron.* **50**(1), 214–224 (2004)
15. Langford, S.K., Hellman, M.E.: Differential-linear cryptanalysis. In: *Annual International Cryptology Conference*, pp. 17–25. Springer (1994)
16. Mahmoud, E.M., Abd, A., Hafez, E., Elgarf, T.A., et al.: Dynamic AES-128 with key-dependent S-box (2013)
17. Manavski, S.A.: CUDA compatible GPU as an efficient hardware accelerator for AES cryptography. In: *ICSPC 2007. IEEE International Conference on Signal Processing and Communications, 2007*, pp. 65–68. IEEE (2007)
18. Matsui, M.: Linear cryptanalysis method for des cipher. In: *Workshop on the Theory and Application of of Cryptographic Techniques*, pp. 386–397. Springer (1993)
19. Matsui, M.: The first experimental cryptanalysis of the data encryption standard. In: *Annual International Cryptology Conference*, pp. 1–11. Springer (1994)
20. Maximov, A.: Some words on cryptanalysis of stream ciphers. Lund Univeristy, Department of Information Technology (2006)
21. Menezes, A.J., Van Oorschot, P.C., Vanstone, S.: Chapter 9: Hash functions and data integrity, pp. 321–383. *Handbook of Applied Cryptography*. CRC Press, Boca Raton, FL (1997)
22. Merkle, R.C., Hellman, M.E.: On the security of multiple encryption. *Commun. ACM* **24**(7), 465–467 (1981)
23. Mohammad, F.Y., Rohiem, A.E., Elbayoumy, A.D.: A novel s-box of AES algorithm using variable mapping technique. In: *Proceedings of the 13th International Conference on Aerospace Sciences and Aviation Technology*, pp. 1–10 (2009)
24. Mukherjee, S., Sahoo, B.: A survey on hardware implementation of idea cryptosystem. *Inf. Secur. J.: Glob. Perspect.* **20**(4–5), 210–218 (2011)
25. Pionteck, T., Staake, T., Stieflmeier, T., Kabulepa, L.D., Glesner, M.: Design of a reconfigurable AES encryption/decryption engine for mobile terminals. In: *ISCAS'04. Proceedings of the 2004 International Symposium on Circuits and Systems*, 2004, vol. 2, pp. II–545. IEEE (2004)
26. PUB, N.F.: 81-des modes of operation (1980)
27. Rapeti, S.A.: NHs: a new non-linear feedback stream cipher. Indian Institute of Technology (2008)
28. Schaefer, E.F.: A simplified data encryption standard algorithm. *Cryptologia* **20**(1), 77–84 (1996)
29. Shannon, C.E.: Communication theory of secrecy systems. *Bell Labs Tech. J.* **28**(4), 656–715 (1949)
30. Shannon, C.E.: Communication theory of secrecy systems. *MD Comput.* **15**(1), 57–64 (1998)
31. Smid, M.E., Branstad, D.K.: Data encryption standard: past and future. *Proc. IEEE* **76**(5), 550–559 (1988)
32. Stalling, W.: *Cryptography and Network Security: Principles and Practices*. Pearson Education India (2006)
33. Stamp, M.: *Information Security: Principles and Practice*. Wiley, New York (2011)
34. Standard, D.E., et al.: Federal information processing standards publication 46. National Bureau of Standards, US Department of Commerce (1977)

35. Van Oorschot, P.C., Wiener, M.J.: A known-plaintext attack on two-key triple encryption. In: Workshop on the Theory and Application of Cryptographic Techniques, pp. 318–325. Springer (1990)
36. Voydock, V.L., Kent, S.T.: Security mechanisms in high-level network protocols. ACM Comput. Surv. (CSUR) **15**(2), 135–171 (1983)
37. Weeks, B., Bean, M., Rozylowicz, T., Ficke, C.: Hardware performance simulations of round 2 advanced encryption standard algorithms. In: AES Candidate Conference, pp. 286–304 (2000)

A Novel Compact UWB Antenna for Wireless Applications



Vishal Pant, Yogendra Choudhary, Abhinav Sharma, Tarun Kumar and R. Gowri

Abstract In this paper, a novel compact monopole microstrip-line-fed Y-shaped patch antenna with an embedded annular ring and L-shaped defective ground plane is presented. The defective ground shape (DGS) ensures that the antenna would work in the ultra-wideband (UWB) (3.1–10.6 GHz). The multiple slots are made in the ground plane which further enhances the return loss (s_{11} parameter) for lower frequencies of the band. The design is printed on FR-4 substrate of 22×24 mm having permittivity $\epsilon_r = 4.4$ and thickness $h = 1.6$ mm. The proposed antenna is designed and analyzed in HFSS software which gives a return loss of 10 dB for the entire ultra-wideband with VSWR lying between 1 and 2 throughout the band. Maximum gain of 5.25 dB is obtained with average being 3.3 dB. With the stable radiation characteristics as described by the radiation patterns of E-plane and H-plane, this antenna will find numerous applications in the new generation of wireless communication systems.

Keywords Microstrip patch antenna · UWB · FR-4 · DGS · VSWR · HFSS

V. Pant · Y. Choudhary · A. Sharma (✉) · T. Kumar · R. Gowri

Department of Electronics Instrumentation & Control Engineering, University of Petroleum and Energy Studies, Dehradun, India

e-mail: abhinav.sharma@ddn.upes.ac.in

V. Pant

e-mail: vishal.pant1@gmail.com

Y. Choudhary

e-mail: 144yogiaict@gmail.com

T. Kumar

e-mail: tkumar@ddn.upes.ac.in

R. Gowri

e-mail: rgowri@ddn.upes.ac.in

1 Introduction

Over the last few years, the subject of microstrip patch antennas [1] has been of immense interest among the academic researchers as well as communication industry. There has been a steep rise in the manufacture of patch antennas which is evident from the recent advances made in the field of wireless communication systems. This has been made possible due to their remarkable ability of being printed directly onto a circuit board. Low cost and ease of fabrication, compact size, wide bandwidth, fewer design parameters, resemblance to proper geometric shapes, compatibility with integrated host system, and unsophisticated feeding techniques are among the primary features of patch antennas [2]. Ultra-wideband is a radio technology that incorporates low energy levels for short-range, high bandwidth communications over a large chunk of microwave spectrum. In 2001, in order to promote very large or wideband transmission bandwidths, the Federal Communications Commission (FCC) authorized the unlicensed band of 3.1–10.6 GHz for commercialized use in UWB applications. This has attracted the attention of several researchers toward UWB technology over the past few years. A major challenge in the design of a UWB antenna is in obtaining optimum results in the entire band prescribed under UWB that is in all the frequencies lying in 3.1–10.6 GHz. This is a cumbersome task as the plot of return loss versus frequency has a tendency to shift toward either lower frequency ranges or higher frequency ranges but to maintain return loss below –10 dB for the entire band requires modifications in the feeding structure and appropriate shape and size of slots carved in the ground plane [3]. Another prominent technique of obtaining desirable results in UWB is as described in [4] which is based on CPW feeding mechanism which offers several advantages over conventional microstrip feed line and significantly reduces the size of the antenna maintaining a wide operating bandwidth. In [5], a complex annular-ring-shaped design is proposed with an emphasis on the DGS approach which leads to remarkable results in both lower and higher frequencies. This is further supplemented by the results obtained in [6]. A comparative analysis of various shapes of patch is done in [7] with tapered patch giving the most desirable results. In [8], a rather unconventional heart-shaped design with modified ground plane is proposed giving a very large operating bandwidth.

In this paper, a compact microstrip-line-fed design of a patch antenna is presented. It consists of a simple Y-shaped structure with an annular ring embedded in it and a separate multi-slotted inverted L-shaped defective ground plane. It is of fairly compact size with overall size being $22 \times 24 \times 1.6 \text{ mm}^3$. A major emphasis has been on improving its bandwidth and the proposed structure achieves good impedance matching with a fractional bandwidth of 111.76%.

2 Antenna Structure and Design

The geometry of the proposed novel UWB antenna is shown in Figs. 1 and 2. The proposed antenna is having the overall dimensions of only 22 mm × 24 mm. The anticipated antenna has two concentric circles of radii R_1 and R_2 , respectively, having a difference of 2 mm are amalgamated with the Y-shaped structure and it is acting as the microstrip feed line of length L_1 and width W_1 . Further, two small circles having radii R_3 and R_4 are placed diagonally to each other and united with the whole patch. These circles are used for the optimization of results in mid-frequency range. Primarily, the ground plane configuration consists of rectangular shape. Then it is further amended into L shape by cutting a rectangular portion of length T_1 and width B_1 . The persistence of L-shaped structure is to accomplish optimum return loss in the lower frequency region and to provide the widespread frequency band. Additional modification in L-shaped ground plane is done through two slots having length B_3 and width T_4 . The purpose of these slots is to proliferate the sharpness of blunt peaks of return loss. The design parameters of proposed patch antenna are shown in Table 1.

Fig. 1 Schematic configuration of the proposed Y-shaped patch antenna with embedded annular ring and L-shaped defective ground plane

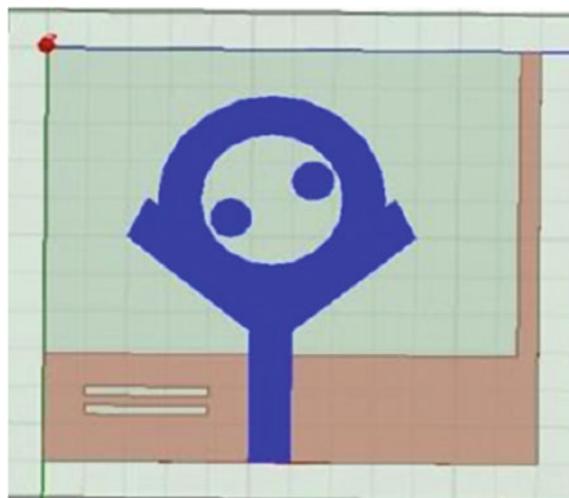


Table 1 Design parameters of patch antenna

Parameters	R1	R2	R3	R4	L1	L2	L3	W1
Unit (mm)	3.5	5.5	1	1	7	7.81	2.236	2
Parameters	T1	T2	T3	T4	B1	B2	B3	B4
Unit (mm)	16.25	22	5.75	0.5	23	24	6	1

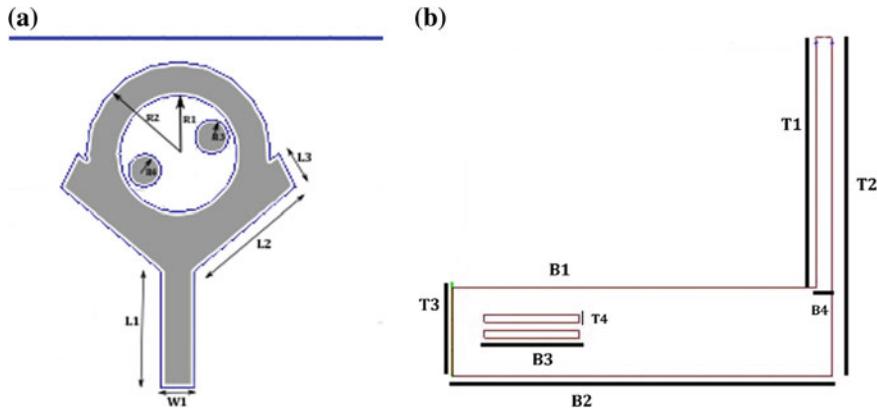


Fig. 2 **a** Top view. **b** Bottom view of the proposed novel compact UWB antenna

3 Results and Discussion

The simulation results were obtained using HFSS software, by Ansoft Corporation which is based on finite element method (FEM) technique. It is clear from Fig. 3 that the return loss is less than 10 dB throughout the ultra-wideband. The lowest dip obtained is -40 dB around 3 GHz. The second dip is at 8.25 GHz and comes out to be approximately -25 dB. Thus, the result is in accordance with the application of the design. It can be observed that the proposed antenna exhibits an ultra-wideband performance from 3 GHz to 10.7 GHz with a fractional bandwidth of 111.76%. The VSWR is depicted in Fig. 4 which varies between 1 and 2 throughout the operating bandwidth. VSWR is a matching coefficient whose value indicates how perfectly an antenna is matched with the transmission line. Therefore, at 3 and 8.5 GHz, VSWR is approximately 1 as perfect matching is achieved at these two resonant frequencies.

The gain versus frequency 2-D plot is shown in Fig. 5a with average gain being around 3.3 dB and peak gain of 4.2 dB at 9 GHz. The graph of efficiency versus frequency is shown in Fig. 5b which indicates that efficiency ranges from 88% to 96% throughout the operating band.

The 3-D polar plot is depicted in Fig. 6 which shows that the maximum gain obtained is around 5.27 dB. The average value is found to be around 3.3 dB. Figure 7a, b shows 2-D radiation pattern in E- and H-plane at two different sampling frequencies, i.e., 3.1 and 8.5 GHz. The simulation results show that nearly good omnidirectional radiation pattern is obtained at these two frequencies and the proposed novel compact antenna is suitable for different wireless applications.

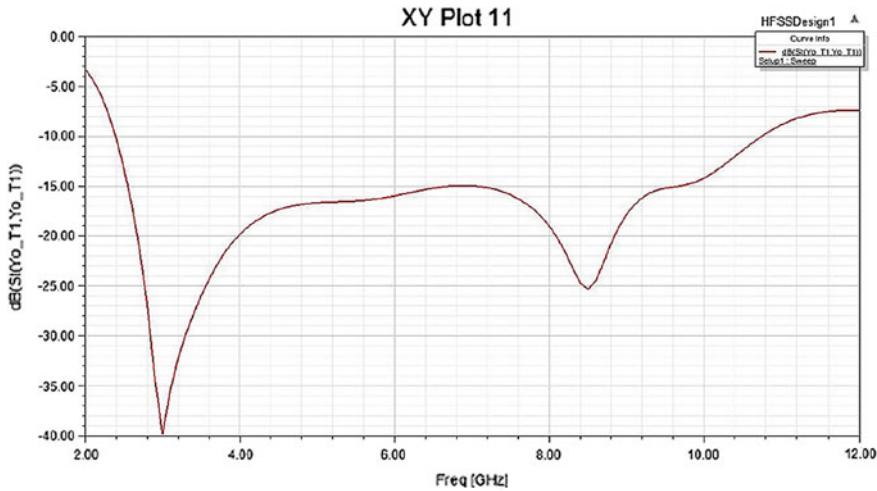


Fig. 3 Simulated results of return loss of the proposed UWB antenna

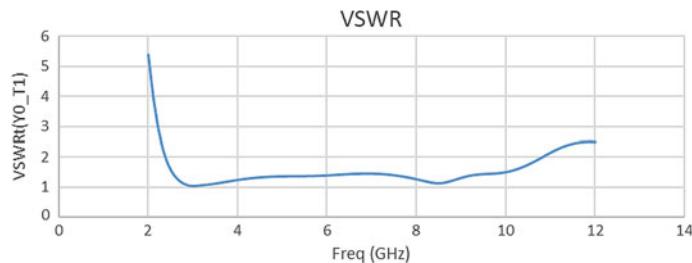


Fig. 4 Simulated results of VSWR of the proposed UWB antenna

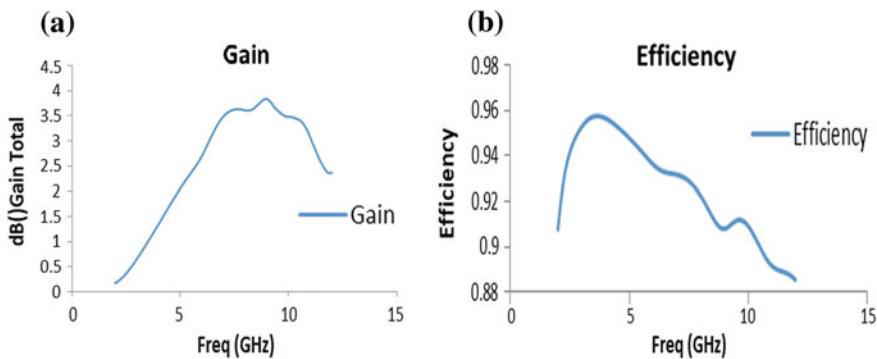


Fig. 5 **a** Simulated gain. **b** Efficiency of the proposed UWB antenna

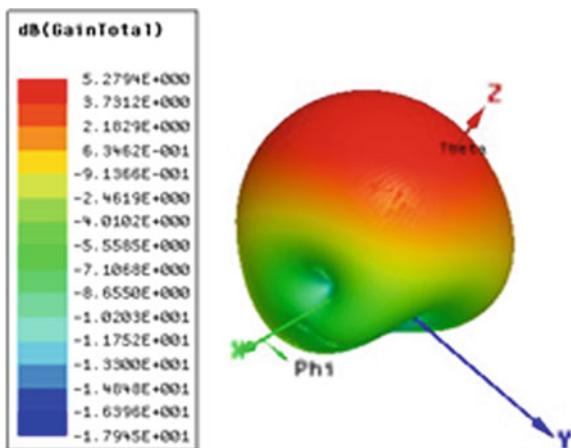


Fig. 6 3-D polar plot along with gain

4 Conclusion

A novel compact design of a microstrip-line-fed patch antenna is successfully implemented. The prototype has dimensions of $22 \times 24 \times 1.6 \text{ mm}^3$ and is fabricated on an inexpensive FR-4 substrate. The simulated results showcase favorable performance in the ultra-wideband. Return loss of 40 dB and -25 dB is obtained at 3 GHz and at 8.25 GHz, respectively, which indicates good impedance matching in both lower as well as higher frequencies. Small size, high efficiency, and stable radiation patterns over the operating bandwidth make this design a viable prospect for UWB applications.

The future scope of this work will involve band notching for frequencies in the range of 3–4 GHz. As the antenna covers the entire ultra-wideband, the band employed by Wi-Max/Wi-Fi will lead to undesirable interference. To avoid this, notching will be done to remove that band such that antenna will no longer produce desired results for those frequencies. In this way, the interference created during the operation of antenna can be eliminated.

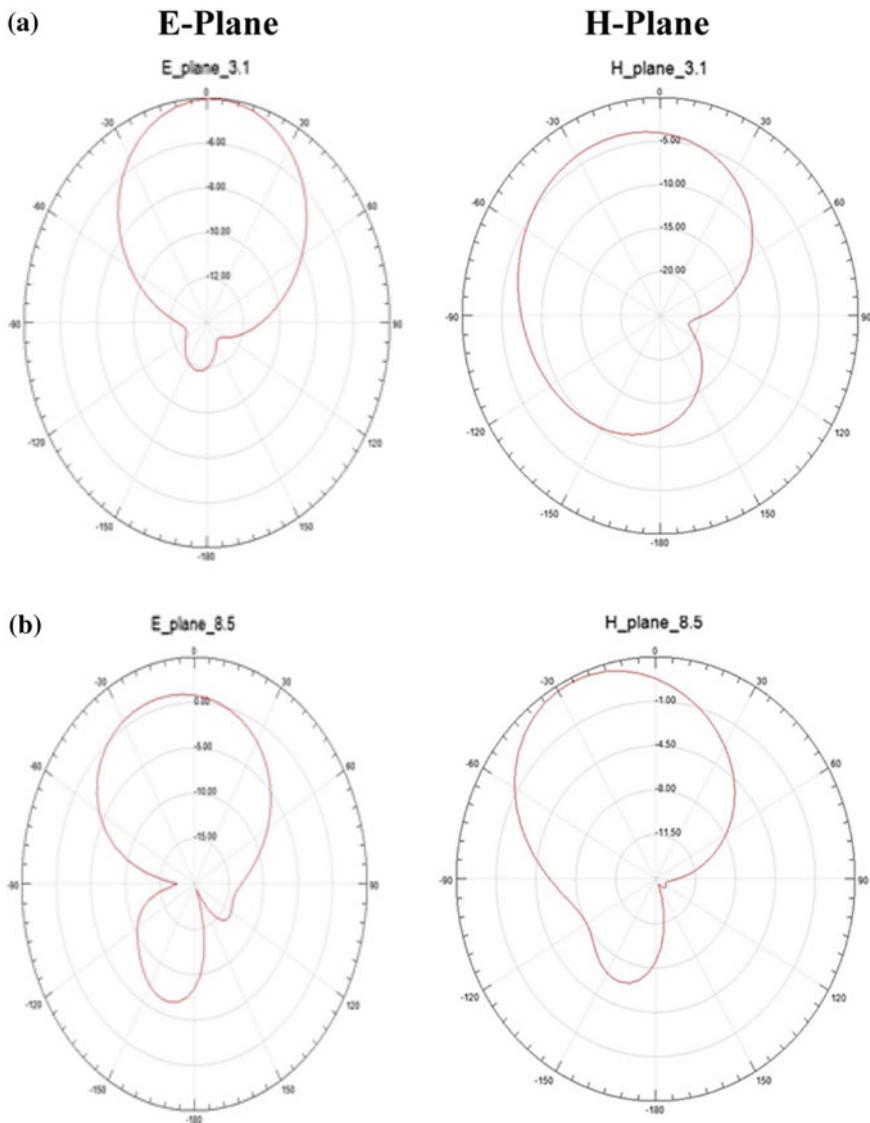


Fig. 7 Radiation pattern at two different sampling frequencies **a** 3.1 GHz **b** 8.5 GHz

References

1. Balanis, C.A.: Microstrip antennas. *Antenna Theory: Analysis and Design*, vol. 3, pp. 811–882 (2005)
2. Eskandari, H., Azarmanesh, M.: A novel broadband design of a printed hexangular slot antenna for wireless applications. *J. Electromagn. Waves Appl.* **22**(8–9), 1273–1282 (2008)
3. Ghaderi, M.R., Mohajeri, F.: A compact hexagonal wide-slot antenna with microstripfed monopole for UWB application. *IEEE Antennas Wirel. Propag. Lett.* **10**, 682–685 (2011)
4. Gautam, A.K., Chandel, R., Kanaujia, B.K.: A CPW-fed hexagonal-shape monopole-like UWB antenna. *Microw. Opt. Technol. Lett.* **55**(11), 2582–2587 (2013)
5. Chandel, R., Gautam, A.K., Kanaujia, B.K.: Annular-ring antenna for UWB applications. In: *Applied Electromagnetics Conference (AEMC)*, pp. 1–2. IEEE Press (2015)
6. Kumar, P., Reddy, G. S., Alex, Z. C.: Planar monopole antenna with defected ground plane for UWB application. *India Conference (INDICON)*, pp. 1–3. IEEE Press (2011)
7. Azim, R., Islam, M.T., Misran, N.: Compact tapered-shape slot antenna for UWB applications. *IEEE Antennas Wirel. Propag. Lett.* **10**, 1190–1193 (2011)
8. Torres, C.A.F., et al.: Heart shaped monopole antenna with defected ground plane for UWB applications. *11th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*, pp. 1–4. IEEE Press (2014)

Monitoring and Control of City Water Distribution System



Shubham Kumar, Nishant Kumar and Arpit Jain

Abstract Water has been one of the most precious things on the earth from the time immemorial. But deforestation, which led to change in climate, has resulted in the decrease in the rate of rainfall. Overexploitation and wastage of water are some of the contributing factors due to which the underground water level is falling very rapidly, and hence the people residing in cities are dependent on water supply and are facing a lot of problem due to mismanagement of water supply. Lack of monitoring leads to wastage and uneven distribution of water. There are many other problems such as leakage in water pipelines. To overcome these problems, there must be a proper and continuous monitoring, controlling, and planned supply of water, so that all the localities and colonies of the city get equal distribution of water, as well as the loss and wastage of water supply should also be controlled. An application-based embedded system is designed with the help of LabVIEW (PLC) and controllers in IOT platform in which by using a controller at main supply station the equal amount of water is regulated and given to all the localities and colonies of the city, and the wastage of water can also be monitored and controlled, and it will not require any manual system so the efficiency and the accuracy will also be more.

Keywords PLC · Arduino · Real-time monitoring · Bluetooth module

1 Introduction

A distribution system can be broadly defined as the methods and facilities through which water is supplied from origin to end users. The main objective of any distribution system is to provide adequate or a sufficient amount of water while maintaining the quality and pressure of water flow. An ideal distribution system should be in such a way that it is capable of providing water to all the areas of the city while maintaining the quality, quantity as well as the pressure of supply. It should also be

S. Kumar · N. Kumar · A. Jain (✉)
UPES, Dehradun, Uttarakhand, India
e-mail: ajain@ddn.upes.ac.in

able to meet the demand of emergency requirement of water like in case of fire; the distribution system should be such that it meets the requirement in that particular area with undue delay and wastage.

Pooja Narkhede et al. designed a PLC-based automatic corporation water-level monitoring and distribution system, which will keep the event of taking and spilling of water using Schneider NG16DL PLC and flow sensors, and the system is divided into three main modules—sensing, decision-making, and implementation. The readings of the sensors are considered by the PLC to take the required decision and this decision is implemented by the PLC through a relay switch [1]. Shaikh Mudassir Nadeem et al. developed a demo model of PLC and SCADA for smart city water supply control system which is going to work remotely and manually using different components like motor, switches, solenoid valve, and sensors [2]. Abhijeet Mancharkar et al. designed a PLC-based water distribution system which will help to improve the performance of the system with less human efforts using 8–10 pressure loggers at various locations. All operations of ESR will be carried out by PLC, which controls inflow and outflow [3]. Mancharkar et al. developed a system using PLC, SCADA framework, solenoid valves, and flow sensors in which the water supply can be controlled by utilizing PLC to decrease wastage of water and handles the immediate control of the physical segments though the control charges are given by the SCADA framework [4]. Rahate1 et al. proposed a method using PLC which is a basic piece and it lessens manual impedance and wastage of water and the level sensor faculties the level of the water and flags the PLC to stop the stream of water through the solenoid valve, in this way limiting the spill out of the fundamental tank [5]. Harish et al. designed a centralized water circulation checking and controlling system using PLC, SCADA, GSM module, and flow sensors in which PLC is used to control the distribution of water and it gives signals to the solenoidal valve according to the input given by GSM module [6]. Prof. Onkare et al. developed a system using PLC, and SCADA is developed to overcome problems like leakage and improper water supply in which automation is done using PLC and SCADA, and both of them will monitor the water distribution system [7]. Aziz et al. designed a controlling and monitoring system using PLC, and SCADA is developed to automate the water level of storage tank. PLC will impart through its information and yield to send the information to the checking framework [8]. Rote et al. developed an automated water distribution system with monitoring of the system is developed using PLC. PLC programming language is done using ladder diagram programming, and it is used to protect the leakage of water and constant pressure maintenance while water distribution [9]. Ejio for Virginia Ebere (Ph.D.) et al. developed a control system using microcontroller which will automate the process of water pumping in an overhead tank storage system which will also detect the water level [10]. Hemrajan et al. used a method by which flow of water in a proper channelized manner is done using embedded technology. They have used a microcontroller to regulate the quantity of water in regular interval of time along with solenoid valve and GSM module to supply water and for wireless communication [11]. Jaiad et al. made a system which is based on IOT using Arduino Nano and Raspberry Pi, and flow sensor is designed to overcome the problem of wastage of water. The Arduino Nano is a microcontroller-based ATMEGA which

will collect data from level sensors [12]. Kumbhar et al. developed an automatic water supply management system is proposed using a microcontroller and GSM module which will efficiently distribute water to all. Arduino microcontroller reads input and display output [13].

The objective of the proposed method is to design a control philosophy which optimizes the water distribution system that should be

- Able to provide an equal amount of water to all the zones of the city and to reduce the wastage.
- Able to detect any type of leakage during channel transmission.
- Able to respond in case of any emergency.
- Generate automatic billing for water usage.
- Block the water supply to the units having unpaid bills.

2 Block Diagram and System Description

A system has been designed for monitoring and control of city water distribution system using PLC, microcontroller, solenoidal valve, Bluetooth module, flow sensors, and LCD (16*2). The framework accordingly comprises fundamental supply framework, what's more, two substations. The principle supply framework is associated with both the substations through pipes. The fundamental supply system is appended to a solenoidal valve which will control the stream of water. If the tank in substation is filled, then it will stop the stream and bad habit vice versa. Three circuits have been composed of utilizing microcontroller and PLC, and the fundamental supply framework is interfaced with a circuit having PLC and microcontroller. The particular substations will be interfaced with a circuit comprising microcontroller, i.e., Arduino UNO. The framework is utilized as a part of two courses—In the main way, the Bluetooth module connected to both the substations will give the message or information to the principle supply and second strategy is that the Bluetooth module associated with primary supply framework is sending the information to the separate substations. The water is supplied from the main station to the substations 1 and 2, respectively. The valves are attached to the main supply system. The PLC controls the opening and closing of the solenoidal valve. Water is supplied from the main system to substation. Ladder logic program of PLC determines the time period for the supply. In our project, we have set a timer of 10 s. As soon as the timer reaches 10 s, the valve which is attached to substation 1 turns off and the valve to substation 2 turns on which leads to the flow of water to substation 2. Microcontroller (Arduino UNO) attached to the supply system and substation communicates with each other by BT module. Arduino and PLC interfacing are done through relay circuit (Fig. 1).

The system includes a main supply system and two substations. The water coming from the main supply station is first passed through the solenoidal valve and after that, it goes to the substation side through pipes. In the main supply system, PLC is used to control the level of water. With the help of ladder programming (PLC), the

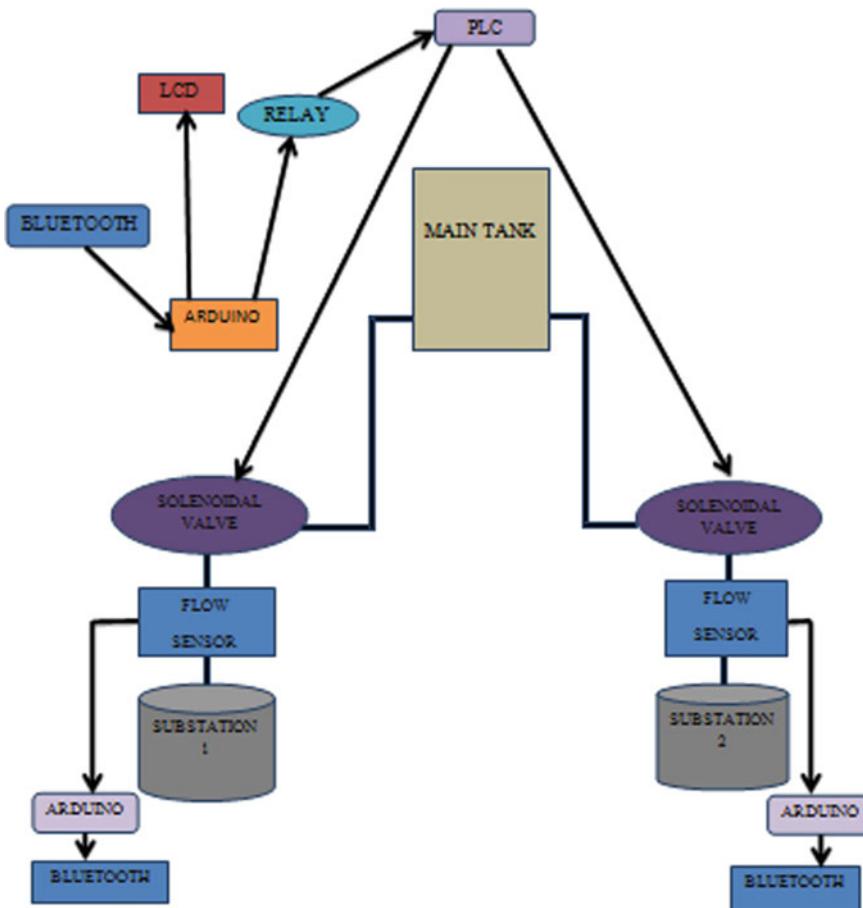


Fig. 1 Block diagram of distribution system

water will go to the first substation for 10s and automatically the flow of water stops and then it will go to the second substation. Further, the master Bluetooth module which is attached in the main supply system gathered the information of the slave module attached to the respective substation. The data about the quantity of water flowing through each substation tank is displayed through LCD (Figs. 2 and 3).

Pseudocode for Main Station:

```

Start
Water sent to sub 1;
Wait(10 sec);
{
Receive (Substation 1);
Wait(10 sec);
  
```

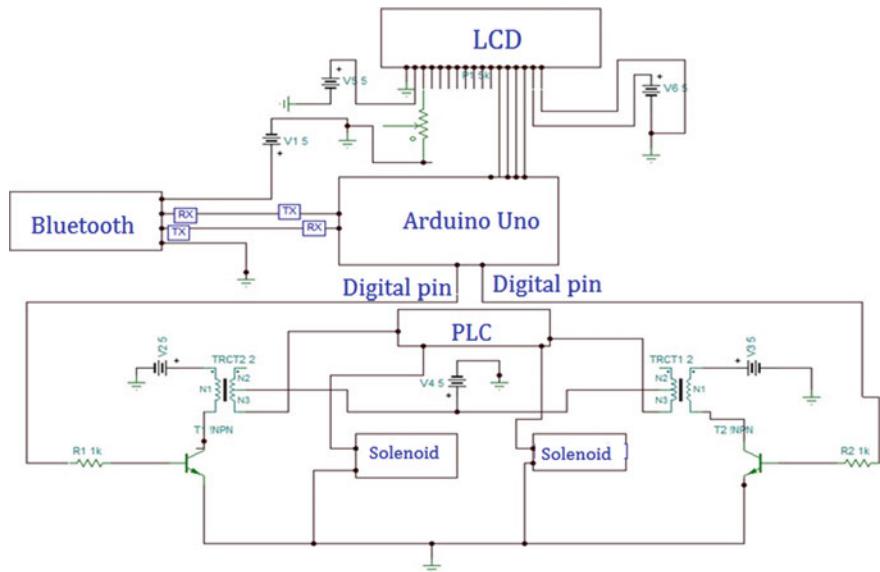


Fig. 2 Circuit diagram of main supply station

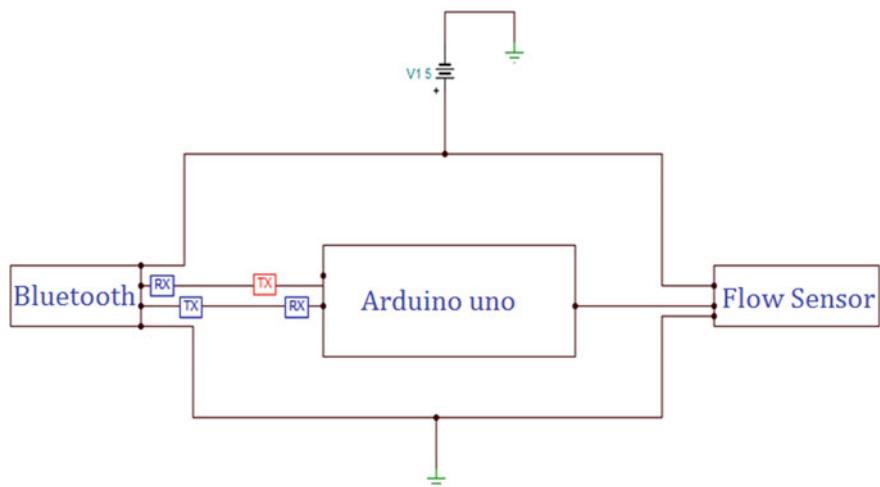


Fig. 3 Circuit diagram of substation

```

Receive (Substation 2);
}
END

```

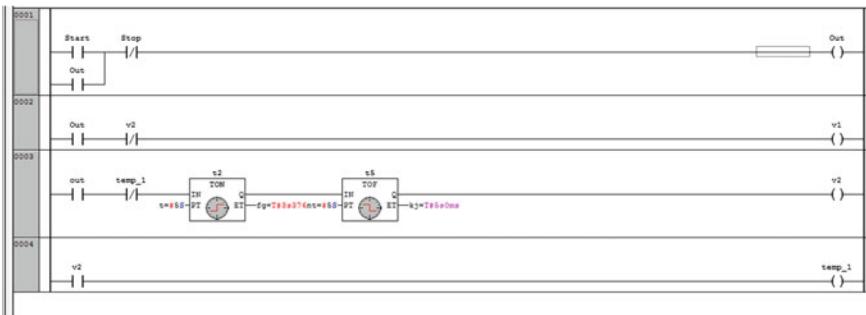
Pseudocode for Substation:

```

Start
Water received at Substation 1;
Wait (10 sec);
{
Send (Main Station);
Wait(10 sec);
}
END

```

3 PLC Program



4 Results and Conclusions

See Table 1.

Table 1 Flow of water in substations (flow rate = 0.001 m³/s)

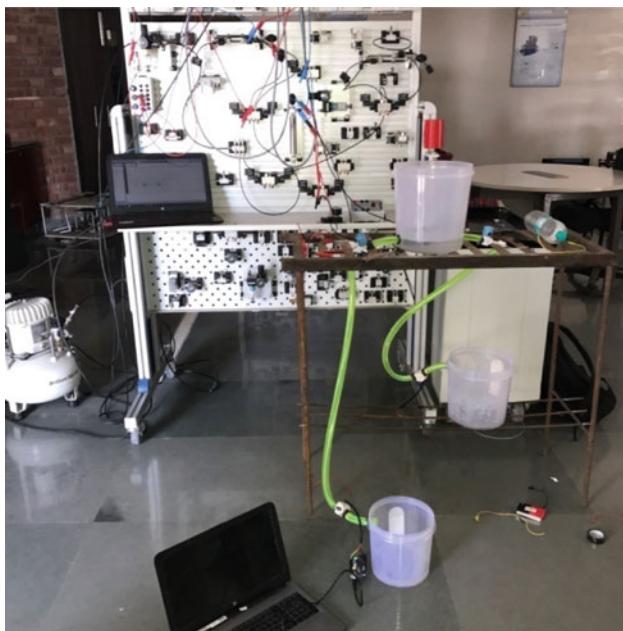
Valve 1 (V1)	Valve 2 (V2)	TIME (s)	Volume of water in substation (in liters)
ON	OFF	2	2
OFF	ON	2	2

5 Hardware Development

A system has been designed for monitoring and control of city water distribution system using PLC, microcontroller, solenoidal valve, Bluetooth module, flow sensors, and LCD (16*2). The framework accordingly comprises fundamental supply framework and two substations. The principle supply framework is associated with both the substations through pipes. The fundamental supply system is appended to a solenoidal valve which will control the stream of water. If the tank in substation is filled, then it will stop the stream and vice versa. Three circuits have been composed of utilizing microcontroller and PLC, and the fundamental supply framework is interfaced with a circuit having PLC and microcontroller. The particular substations will be interfaced with a circuit comprising microcontroller (Arduino UNO). The framework is utilized as a part of two courses—In the main way, the Bluetooth module connected to both the substations will give the message or information to the principle supply and the second strategy is that the Bluetooth module associated with primary supply framework is sending the information to the separate substations. Thus, broad development of populace improvement and innovation has prompted the need for appropriate use of the common assets particularly water. Consequently, our proposed framework and the audit of all the conceivable execution of innovation is the initial move toward aversion and legitimate usage of water. The audit of mechanized water dissemination framework with the different controllers and parameters concentrates on the elements, for example, appropriate supply, red caution pop-ups, filtration, stream control, and supervision utilizing different conventions, is closed with the future parts of constant execution in the metropolitan organizations where a shortage of water is the gigantic issue.

The future scope of the project is as follows:

1. This system can be implemented at the industrial level using SCADA.
2. By installing sensors on pipelines, the exact location of leakage can be known.
3. Water contamination can also be detected by a flow sensor which determines the volume of flow at any instant.
4. Proper distribution of water can be ensured for a different population.



Working model of distribution system

References

1. Narkhede, P., Bholane, A., Mirza, R., Jawale, P.: Electrical Engineering(Electronics and power), PLITMS Buldana(India) National Conference CONVERGENCE 2016, 06–07th April 2016
2. Baranidharan, T. (Associate professor), Chinnadurai, A., Gowri, R.M., Karthikeyan, J.: (UG students, EIE, K.S. Rangasamy College of Technology) in Int. J. Electr. Electron. Eng. (IJEEE), 7(1), (2015)
3. Nadeem, S.M., Imtiyaz, M.A., Patel, P.R., Ali, S.Z., Prof. Losarwar, V.: BE Student, Asst. Professor, P.E.S. College of Engineering, Aurangabad, Maharashtra, India, vol 2 issue 01 April 2016
4. Hattikatti, P., Karwande, S., Bhandarkar, P.S.R.M.: UG Student, Department of E&TC, PCCOE, Pune, Maharashtra, India 1,2,3 A.P. Department of E&TC, PCCOE, Pune, Maharashtra, India vol 5, issue 6, June 2017
5. Mancharkar, A., Kulthe, R., Shewale, I.: Department of Electronics & Telecommunication Sandip Institute of Technology & Research Centre, Mahiravni, International Journal of Emerging Technologies and Engineering (IJETE), vol. 3 Issue 3, March 2016, ISSN 2348–8050
6. Rahate1, A., Gaikwad2, A., Rehu 3, D., Raichurkar4, A., Jadhav5, R.: B.E Students1,2,3,4, Assistant Professor 5 Department of Electronics Engineering, V.E.S.I.T, International Journal of Advance Engineering and Research Development Volume 4, Issue 4, April 2017
7. Harish, K.M., Chaitra, R., Deepika, R., Divya, K.S., Nandini, M.: Int. J. Sci. Eng. Technol. Res. (IJSETR) 5(6) (2016)
8. Prof. Onkare1, R.P., Randhir Mane2, S., Rakshe3, P.R., Waghmare4, P.S.: Assistant Professor, Department of Electronics and Telecommunication Engineering, P.V.P. Institute of Technology, Budhgaon, Maharashtra, India, vol. 5, Issue 3, March 2017

9. Aziz, S.A.B.A.: Faculty of Electrical and Electronic Engineering University Tun Hussein Onn Malaysia July 2013
10. Rote, S., Mourya, A., Oak, T.: Electronics and Telecommunication, Prof. Abhimanyu Yadav A.P. Department of E&TC, SITRC, Nashik, India, vol. 6 Issue 04, April-2017
11. Ebere, E.V., (PhD), Francisca, O.O. (PhD): Lecturer, Department of Computer Science Nnamdi Azikiwe University Awka, Nigeria, vol. 1, issue 6, August 2013
12. Hemrajan, A.R., Patel, B., Asst. Prof. Sheth, S.: 2 Student, Electronics and Communication Engineering, ITM Universe, Vadodara, (RTEECE 08th–09th April 2016
13. Jaiad *1, A.T., Ghayyib 2*1, H.S.: The Ministry of Electricity of Iraq, Electricity Distribution Directorate of Babylon, Babylon, Iraq2 Department of Computer Engineering Techniques, The Islamic University College, Babylon, Iraq [Jaiad et al., vol.5 (Iss.5): May, 2017]
14. Bardavelidze1, A., Basheleishvili2, I., Bardavelidze31 Akaki, K.: Tsereteli State University, Kutaisi, Georgia 2Akaki Tsereteli State University, Kutaisi, Georgia 3Georgian Technical University, Tbilisi, Georgia

Detection of DTMF by Using Goertzel Algorithm and Optimized Resource-Sharing Approach



B. Khaleelu Rehman, Adesh Kumar, Salauddin Mohammad,
Mudasar Basha and K. Venkata Siva Reddy

Abstract Dual-Tone Multi-frequency (DTMF) is the Standardized term in the telecom industry, where if anyone press a key two tones are generated simultaneously. The DTMF tone detection is an important block in several embedded applications, which works for low-power devices. The DTMF detection using high-end FPGA is carried out by using the Goertzel algorithm and further the design is optimized by using the resource-sharing approach. The resource-sharing approach uses less area, less delay, and low power. Very few multiplexers, adders are used as compared to the conventional approach of DTMF detection. All the timing diagrams are analyzed using the ModelSim simulation software. Pre-synthesis and post-synthesis are carried out using the XILINX Vivado. Virtex-5 FPGA device is used to check the functionality of the hardware.

Keywords FPGA · Xilinx Vivado. Virtex-5 · ModelSim · DTMF

B. K. Rehman (✉)

Department of Electrical & Electronics Engineering, University of Petroleum & Energy Studies (UPES), Dehradun, Uttarakhand, India
e-mail: krehman@ddn.upes.ac.in

A. Kumar

Department of Electrical and Electronics Engineering, School of Engineering, University of Petroleum and Energy Studies, Dehradun, India

S. Mohammad

JB Institute of Engineering & Technology, Hyderabad, Telangana, India

M. Basha

BVRIT, Hyderabad, Telangana, India

K. Venkata Siva Reddy

Ravindra College of Engineering for Women, Kurnool, Andhra Pradesh, India

1 Introduction

The applications of the DTMF [1] are in the credit inspection, telephone/mobile phone touch-tone phone dialing, and so on. The Goertzel algorithm is used for the DTMF generation and detection. From the Fig. 1 the DTMF signal sends two frequencies at the same time [2–4] one from the low-frequency group and another from the high-frequency group. For example, if key ‘5’ is pressed then the frequencies 770 Hz from the low-frequency group and 1336 from the high-frequency group is generated.

2 Literature Review

Detection of the DTMF by using the Fast Fourier Transform (FFT) is one approach but the problem for the FFT technique is taking more area on the hardware usage, i.e. FPGA. Shaterian and Gharaee [5] used Goertzel algorithm using a resource sharing approach. Xingi [2] has proposed and compared the modified Goertzel to the conventional Goertzel algorithm approach for the DTMF detection. Dubey et al. [6] proposed a WSN (Wireless sensor network) and the embedded systems approach for the DTMF detection for the application of the irrigation system. Niheai et al. [7] proposed a cargo drop approach for the DTMF detection through the adaptive omission for the frequency analysis and shows the 50% reduction in the load. The problem from the above literature review shows the DTMF detection is possible with the separate ASIC and used the complex algorithm. Fast Fourier transform is one technique but it consumes high power and large area. The proposed solution is the efficient area by using the split Goertzel algorithm and further reduced by using the sharing of the FPGA resources approach.

Fig. 1 DTMF keypad low- and high-frequency group

		High Frequency group				
		1209 Hz	1336 Hz	1477 Hz	1633 Hz	
Low Frequency group		697 Hz	1	2	3	A
		770 Hz	4	5	6	B
		852 Hz	7	8	9	C
		941 Hz	*	0	#	D

3 Proposed Method

Block diagram for the DTMF detection is shown in Fig. 2. The 4*4 hex keypad is connected physically to the FPGA board. In the absence of the hex keypad, the slide switches in the FPGA acts as the input. The signal generator block is for the frequency testing block, which includes the DDS core and the tone generator. The DDS core [2] is an inbuilt analog input signal available in the Xilinx IP core. The tone generator block combines the two frequencies into one. The next module is the noise-added module. The AWGN [8] is added with the tone generator. Additive White Gaussian noise input consists of the noisy signal and that signal is given as the input to the frequency detection block module.

The frequency detection module for the DTMF decoder can be done with the following solutions:

- (1) 256-point FFT
- (2) Goertzel algorithm
- (3) Resource-sharing approach.

3.1 256-Point FFT

The 256-point Fast Fourier Transform can be calculated by using the Xilinx IP core available in the Xilinx Vivado software. 256 point DIF-FFT is used for finding the magnitude and phase response. The FFT algorithm calculates eventually across the entire bandwidth of the incoming signal. The block diagram shown in Fig. 2 has the frequency detection module block. That block is replaced with the Goertzel algorithm for the detection of the DTMF signal.

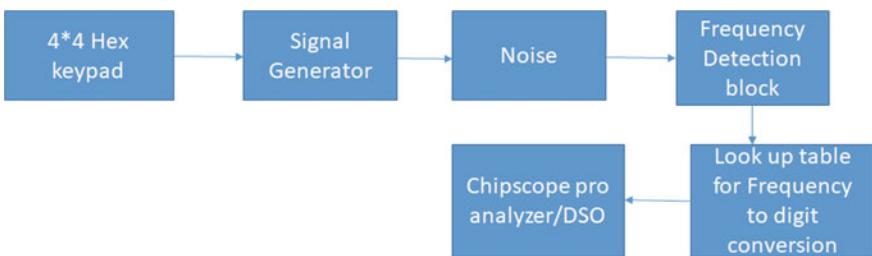


Fig. 2 Block diagram for DTMF detection

3.2 Goertzel Algorithm

The Goertzel algorithm identifies the frequency components of the signal. Goertzel algorithm is used to compute specific predetermined frequency. The second-order IIR filter is used to remove the frequency components and the transfer function of the Goertzel algorithm equation is shown in Eq. (1).

$$H(Z) = \frac{1 - W_N^K Z^{-1}}{1 - (2 \cos \frac{2\pi k}{N})Z^{-1} + Z^{-2}} \quad \text{Where } K = 0, 1, 2, \dots, N-1 \quad (1)$$

The implementation of the Goertzel algorithm by using direct form-II approach is shown in Fig. 3. The required adders are 3 and three multipliers are required for the DTMF detection.

3.3 Resource-Sharing Approach

The resource-sharing approach uses the less number of multipliers and adders as compared to the Goertzel algorithm approach. The memory element (Z^{-1}) in Fig. 3 is replaced with the buffer registers and the multiplexer. The architecture of the resource-sharing approach is shown in Fig. 4. In the Direct form II structure, there are two memory elements and those two memory elements are replaced with Fig. 4 for further reducing the hardware.

The output of the frequency detection module is given as the input to the next module, i.e., frequency to digit converter. The peak magnitude of the higher group and the peak magnitude of the lower group is considered. The chip scope pro analyzer is the software which allows the on-chip verification for the inside the FPGA to validate the result.

Fig. 3 Direct form-II Goertzel algorithm

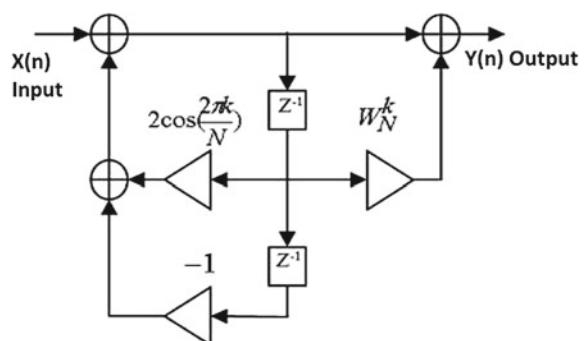
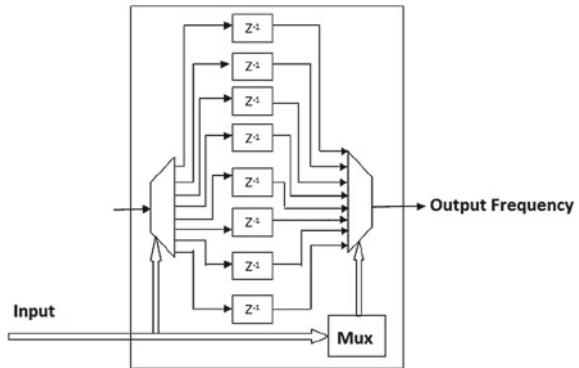


Fig. 4 Resource-sharing approach Goertzel algorithm



4 Results and Discussion

The ModelSim simulation results of different keys are shown in Fig. 5. The ChipScope results are showing in Fig. 6 (Table 1).

The results listed in Table 1 presents the synthesis report generated in the Xilinx Vivado software with number of slices, look-up tables, flipflop pairs, number of input–output block, buffer memory and digital signal processing elements, configured for Virtex-5 FPGA hardware.

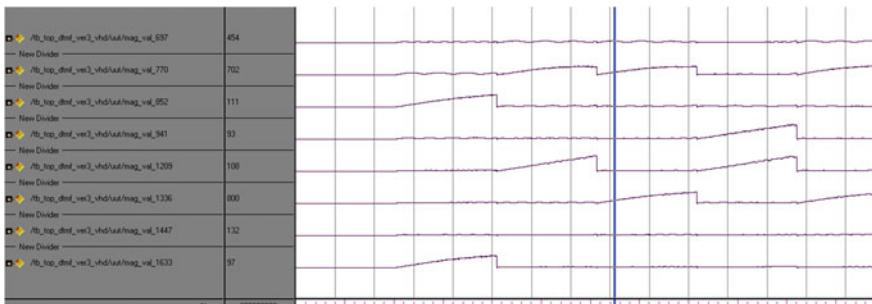


Fig. 5 ModelSim simulated result of DTMF detection

Table 1 Hardware utilization detail on Virtex-5 XC5VX-110t-1ff1136 FPGA

Hardware parameter	256 point FFT	Goertzel algorithm	Resource sharing
Number of slices	6123	3214	1215
Number of slice LUTs	3502	2405	1908
Number of fully used LUT-FF pairs	856	547	386
Number of bonded IOBs	19	19	19
Number of BUFG/BUFGCTRLs	17	8	11
Number of DSP48Es	23	6	3

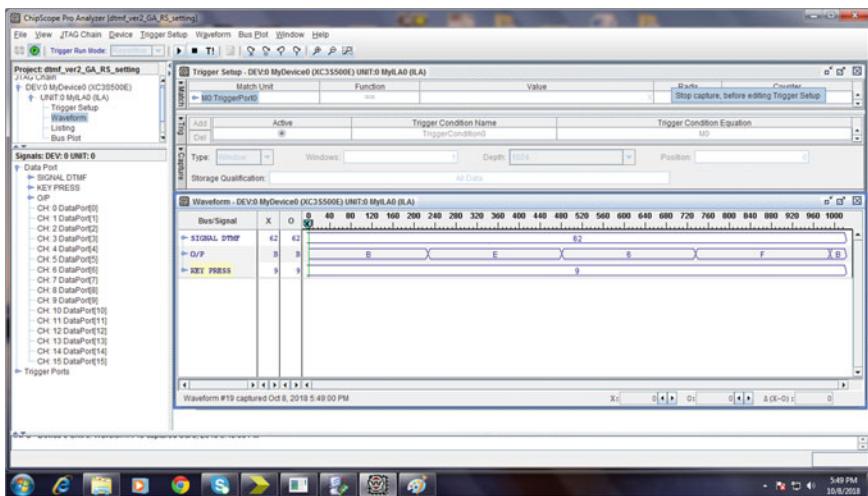


Fig. 6 ChipScope result of DTMF detection using Goertzel and resource-sharing approach

5 Conclusion

The hardware chip of DTMF Goertzel approach is design in Xilinx Vivado successfully and synthesized on Virtex-5 FPGA.

References

1. Vazquez, D., Huertas, G., Avedillo, M.J., Quintana, J.M., Rueda, A., Huertas, J.L.: LP-LV high-performance monolithic DTMF receiver with on-chip test facilities. In: VLSI Circuits and Systems, vol. 5117, pp. 298–310. International Society for Optics and Photonics (2003)
 2. Xinyi, Z.: The FPGA implementation of modified Goertzel algorithm for DTMF signal detection. In: 2010 International Conference on Electrical and Control Engineering (ICECE), pp. 4811–4815. IEEE (2010)
 3. Kherde, S.N., & Hasamnis, M.: Efficient design and implementation of FFT. In: International Journal of Engineering Science and Technology (IJEST), ISSN, 0975-5462 (2011)
 4. Vazquez, D., Avedillo, M. J., Huertas, G., Quintana, J. M., Pauritsch, M., Rueda, A., Huertas, J.L.: A low-voltage low-power high performance fully integrated DTMF receiver. In: Proceedings 27'h European Solid-State Circuits Conference ESSCIRC 2001, pp. 18–20 (2001)
 5. Shaterian, K., Gharaee, H.: DTMF detection with Goertzel algorithm using FPGA, a resource sharing approach. In: 2010 Intl Conference on Electronic Devices, Systems and Applications (ICEDSA), pp. 196–199. IEEE (2010)
 6. Dubey, V., Dubey, N., Chouhan, S.S.: Wireless sensor network based remote irrigation control system and automation using DTMF code. In: 2011 International Conference on Communication Systems and Network Technologies (CSNT), (pp. 34–37). IEEE (2011)
 7. Nihei, K., Nogaki, S., Satoda, K.: Load reduction methods for DTMF decoders with adaptive omission of frequency analysis. In: 2013 IEEE 9th International Colloquium on Signal Processing and its Applications (CSPA), pp. 155–160. IEEE (2013)
 8. www.xilinx.com/support/documentation/user_guides/ug384.pdf

Combined Markov Model and Zero Watermarking for Integrity Verification of English Text Documents



Fatek Saeed and Anurag Dixit

Abstract To assure the integrity of the content by detecting attempted tampering of the original document is integrity verification attempt. The documents are usually watermarked with a watermarking technique. The tampering in document can alter or destruct the watermark and can misuse the document by the unauthorized user. Zero watermarking provides the most accurate results for tampering detection and integrity verification. The watermarking technique has some stages including watermark embedding, watermark extraction, watermark detection, and possible attacks during the watermarking process. The watermark embedding process performed by the embedding algorithm and the watermark extraction process extracts the watermark from the text by extraction algorithm. Combined Markov model and zero watermarking for integrity verification of English text documents are presented in this paper. The letter level- and word level-based combined Markov model and zero watermarking method are analyzed. 3 g LNMZW and 5 g WNMZW algorithms are developed. Based on the N-gram orders of Markov model in word mechanism, each word of the text is considered as 1 g or one order of Markov model based on this procedure the 5 g of Markov model is improved and compared its performance with other algorithms with different attacks in this paper.

Keywords Watermarking · Tamper detection · Markov chain

1 Introduction

The rapid growth of advance information technology has developed and made the digital access as simple. The security of intellectual property rights of digital media has become an important issue due to this growth. The watermarking has obtained widespread attention, and it has many applications such as publication protection, documentation, confidential communication, and measurement [1]. The

F. Saeed (✉) · A. Dixit
Computer Science and Engineering, Galgotias University, New Delhi, India
e-mail: fateksaeed@gmail.com

recent improvements in information and communication technologies gave numerous growths and simplified the digital content distribution, communication, and reproduction. However, some issues and threats are also occurred [2].

The contents of all digital multimedia on the Internet can be categorized into audio and video, with the intention being to achieve protection and feasible communications for different media. The problem of accomplishing the authenticity and integrity verification for text documents were done by zero watermarking techniques. Due to the high complexities in most of the studies on copyright protection for authentication and integrity verification ignored the performance impacts. In some other cases, such methods involved the overhead need for distributing algorithms and/or keys between communicating parties [3].

Most of the watermarking techniques alter the original data during the watermark embedding stage regardless of whether those techniques are in spatial domain or frequency domain. In the purpose of authority protection, the confidential data are embedded but they deform the original data at the same time. This situation represents a challenge between hiding and strength. All what must be done to get rid of this conflict, the watermark must be designed based on the essential letters of the text but not to alter the information or text which is in the source. These kinds of handling which do not make any change in original data are called zero watermarking.

The zero watermarking provides a solution to address the difficulty between invisibility and robustness. Recently, a new technique of zero watermarking based on DWT and Markov model was evolved [4].

2 Combined Zero Watermarking and Markov Model

In this paper, we combine zero watermarking and Markov model; these algorithms are named as LNMZW3 and WNMZW4. In LNMZW3 algorithm, the original script file remains the same for embedding watermark, its logical way. The LNMZW3 algorithm utilizes the Markov model of the natural language in this; Markov models are utilized to check the contents of English script file and detect the possible characteristics of mutual relations among these contents as probabilistic models based on letter method and third order of Markov model which are used to produce the watermark key that is embedded logically within the original script file or saved in the watermark database. These steps are described in Fig. 1.

This section describes how to propose model used English text LNMZW and WNMZ algorithm with N-gram order of Markov model. For example, consider the following sentence:

Each series of three different successive characters in the given text is a state by itself, when we are using 3 g of LNMZW. The procedure of Markov transitions from state to state in given script (Fig. 2).

Thus, the same case for transitions showed in Table 1 when we use 5 g of Markov model to analyze English text documents.

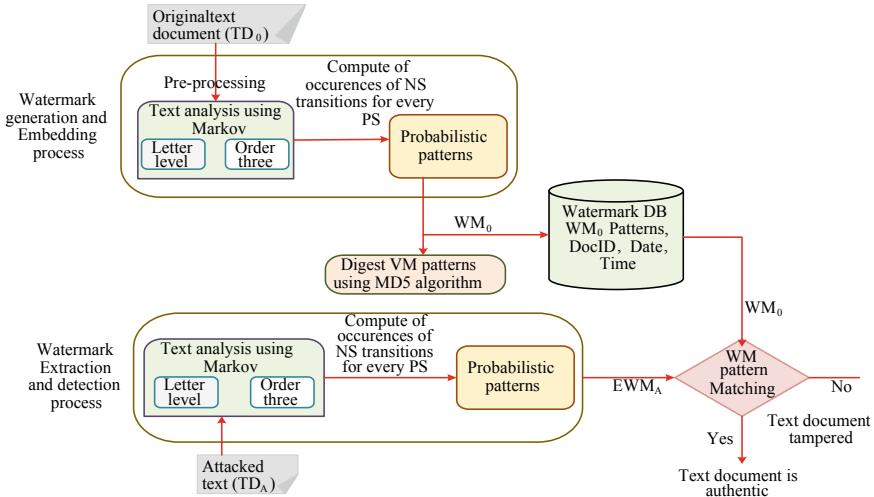


Fig. 1 Markov model-based watermark generation and detection process

"The quick brown fox jumps over the brown fox who is slow jumps over the brown fox who is dead."

Fig. 2 Sample text

"the quick brown fox jumps", "quick brown fox jumps over", "brown fox jumps over the", "fox jumps over the brown", "jumps over the brown fox", "over the brown fox who", "the brown fox who is", "brown fox who is slow", "fox who is slow jumps", ..., "brown fox who is dead".

As far as this script is handled, and if the Markov model is presently at "the quick brown fox" state, the likely transitions for each state that could be:

"the quick brown fox jumps → over", "the quick brown fox jumps → the", "the quick brown fox jumps → brown", "the quick brown fox jumps → fox", ..., "the quick brown fox jumps → dead".

Based on the 5 g order of Markov model on word level to check the sentence, Table 1 is obtained.

From Table 1, if we consider the state "the brown fox who is", the next state transitions are "dead" and "slow".

2.1 Watermark Patterns Generation

According to watermark algorithm rules and preprocessing, we need the original text as input to change uppercase letters into lowercase letters and to delete all spaces from input data, so watermark pattern is produced as an output. This watermark

Table 1 Sample states and transitions of WNMZW5

State ID	State	Transitions
1	(“the quick brown fox jumps”)	[“over”]
2	(“quick brown fox jumps over”)	[“the”]
3	(“brown fox jumps over the”)	[“brown”]
4	(“fox jumps over the brown”)	[“fox”]
5	(“jumps over the brown fox”)	[“who”, “who”]
6	(“over the brown fox who”)	[“is”, “is”]
7	(“the brown fox who is”)	[“dead”, “slow”]
8	(“brown fox who is slow”)	[“jumps”]
9	(“fox who is slow jumps”)	[“over”]
10	(“who is slow jumps over”)	[“the”]
11	(“is slow jumps over the”)	[“brown”]
12	(“slow jumps over the brown”)	[“fox”]

model will be saved in the watermark database with input script file, file identity, author name, current date, and time.

A matrix of transition probabilities can be calculated by Eq. (1).

$$M[ps][ns] = P[i][j], \text{ for } i, j = 1, 2, \dots, n \quad (1)$$

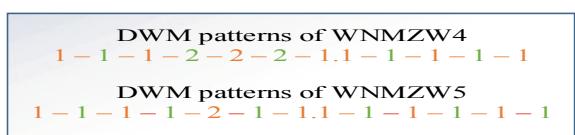
where

- n is the total number of states
- i denotes PS “the present state”
- j denotes NS “the next state”

$P[i, j]$ represents the probability of making a transition from letter set i to letter j .

This procedure obtains all nonzero values of each state from the Markov chain matrix and concatenates them sequentially to generate the original watermark patterns based on WNMZW approach with 4 and 5 g orders presented in Fig. 3.

An MD5 message summary is produced for getting a protected and compact form of the watermark after watermark production as successive patterns, notational as given by Eq. (2)

Fig. 3 The original watermark models of the previous script trial based on WNMZW4 and WNMZW5

$$WMP = MD5(WMP_O) \quad (2)$$

Algorithm1: Watermark generation and embedding algorithm

```

Watermark generation
- Input: matrix of the transition probability from  $i^{th}$  word sets to  $j^{th}$  words ( $MP[ps][ns]$ )
     $WMP_O$ , Digests watermark  $WM_O$ 
- output:  $WMP_O$ ,  $WM_O$ 
-BEGIN
- // Analysis of the given text by computing the total frequencies of the transitions for every state
- Call text_analysis ( $T_p, Ngram$ )
- // concatenate watermark patterns of every states shown in Markov chain matrix
- Loop  $cs = 1$  to  $MP[states].Length$ 
Loop  $ns = 1$  to  $MP[transitions].Length$ 
IF  $MP[cs][ns] = 0$  // states that have non zero transitions
 $WMP_O &= MP[cs][ns]$ 
 $ns ++$ 
- store  $WMP_O$  in DWM database
- // Digest the original watermark using MD5 algorithm
 $WM_O = MD5(WMP_O)$ 
- Output  $WMP_O, WM_O$ 
-END

```

Here, WM_O refers to the original watermark, WMP_O refers to original watermark patterns, and MD5 refers to hash function.

2.2 Experimental Setup and Results

The precision of the tampering detection is calculated. Then the calculated results are compared with the existing previous arrangements of letter level and word level of Markov model on the basis of authentication characteristics, and under the most general nature and volume of possible tampering attacks including insertion, deletion, and reorder attacks in several irregular places of the experimental datasets. This part describes the results and calculation of the LNMZW and WNMZW algorithms.

The performance evaluation of the LNMZW and WNMZW algorithms at various tampering attacks on different dataset sizes was performed for various situations of the attack volumes and types. Then, the performance average was detected and it was compared to present existing algorithms which are LNMZW1, LNMZW2, and WNMZW4. Later level, the great work of which algorithm was discovered attacks, which are intercalation, deletion, and reorder attacks in several random places of the empirical datasets [5] (Fig. 4).

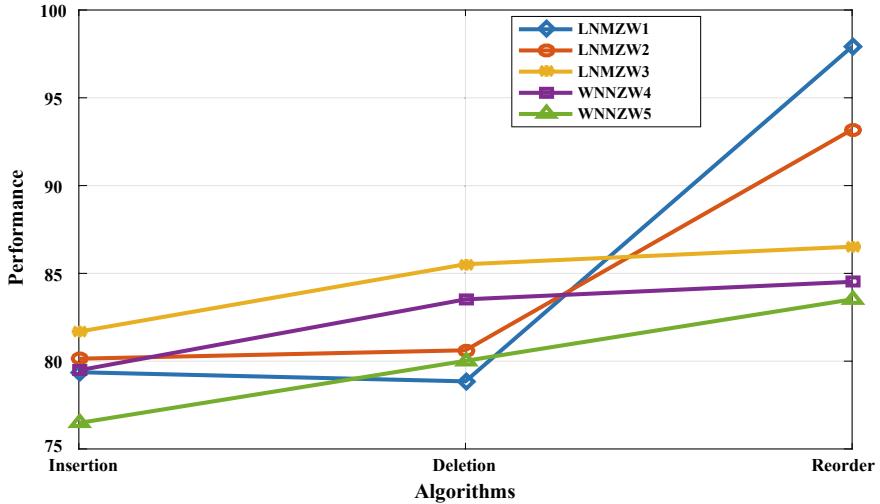


Fig. 4 Performance average of LNMZW1, LNMZW2, LNMZW3, WNMZW4, and WNMZW5 algorithms under all attacks. Figure 3 shows the performance averages of existing algorithms versus all dataset volumes and scenarios of rearranging with all attack volumes

Figure 5 shows the effect of dataset size on watermark robustness under all volumes of all attacks (insertion, deletion, and reordering). Generally, the outcomes clarify that the robustness value got better when Ngram order of Markov model is reduced and vice versa when Ngram order is increased. Based on the comparison, the impact on dataset size on watermark robustness in case of our WNMZW approach is improved when the size of the text document increases. The watermark robustness

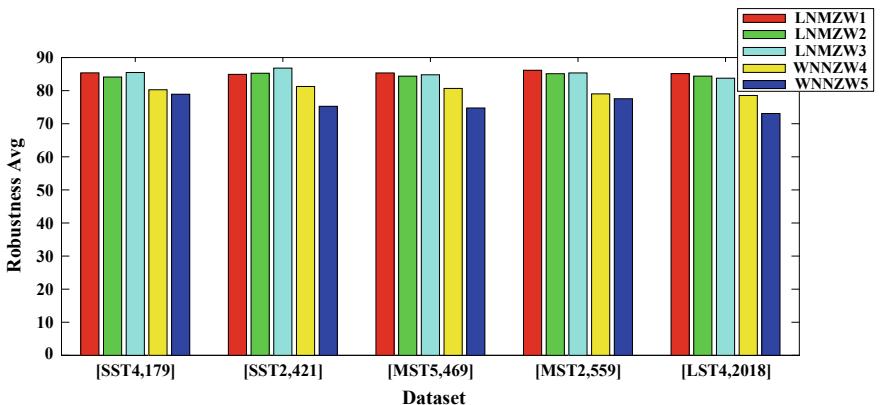


Fig. 5 Comparison of dataset size effect on watermark robustness under all volumes of all attacks

is improved when the size of the text document decreases in the case of LNMZW approach.

3 Related Work

The rapid development of the Internet and availability of networked computers have made the distribution of multimedia data very fast and convenient without losing information. The consequences of such applications lead to modification and distribution of illegal data easier for the unauthorized parties. To overcome these problems, the digital watermarking technique came into existence.

4 Conclusion

In this paper, the Markov model-based watermarking is described. In order to analyze the English script file and discover the mutual relations of contents of the text documents corresponding to the probabilistic patterns of states and transitions finally producing, the third arrangement of LNMZW and fifth arrangement of WNMZW are used. In this study, the LNMZW and WNMZW algorithms are compared with existing LNMZW1, LNMZW2, and WNMZW4 approaches against different datasets.

References

1. Wang, Wei, et al.: A novel robust zero watermarking scheme based on DWT and SVD. In: 2011 4th International Congress on Image and Signal Processing (CISP), vol. 2. IEEE, (2011)
2. Khan, A., Mirza, A.M., Majid, A.: Optimizing perceptual shaping of a digital watermark using genetic programming. Iran. J. Electr. Comput. Eng. **3**(2), 144 (2004)
3. Tayan, O., Kabir, M.N., Alginahie, Y.M.: A hybrid digital-signature and zero-watermarking approach for authentication and protection of sensitive electronic documents. Sci. World J. **2014** (2014)
4. Cao, H., et al.: A zero-watermarking algorithm based on DWT and chaotic modulation. In: Independent Component Analyses, Wavelets, Unsupervised Smart Sensors, and Neural Networks IV, vol. 6247. International Society for Optics and Photonics (2006)
5. Ghilan, M.M., Ba-Alwi, F.M., Al-Wesabi, F.N.: Combined Markov model and zero watermarking techniques to enhance content authentication of English text documents. Int. J. Comput. Linguist. Res. **5**(1), 26–42 (2014)
6. Yang, Y., et al. A novel robust zero-watermarking scheme based on discrete wavelet transform. J. Multimed. **7**(4) (2012)
7. Jalil, Z., et al.: Improved zero text watermarking algorithm against meaning preserving attacks. World Acad. Sci. Eng. Technol. **46**, 592–596 (2010)
8. Kaur, S., Babbar, G.: A Zero-Watermarking algorithm on multiple occurrences of letters for text tampering detection. Int. J. Comput. Sci. Eng. **5**(5), 294 (2013)

9. Ba-Alwi, F.M., Ghilan, M.M., Al-Wesabi, F.N.: Content authentication of English text via internet using zero watermarking technique and Markov model. *Int. J. Appl. Inf. Syst. (IJAIS)* **7**(1), 25–36 (2014)
10. Chroni, M., Nikolopoulos, S.D.: Watermarking PDF documents using various representations of self-inverting permutations. [arXiv:1501.02686](https://arxiv.org/abs/1501.02686) (2015)

Polymer Nanocomposite Membranes for Water Treatment



Madhu Sharma and Anupam Sharma

Abstract Water treatment is a procedure that makes water more sustainable for several residential and commercial purposes through advanced technologies that remove pollutants and contaminants, thereby improving the quality of water. For the treatment of wastewater, membrane technology is quite impressive for a specific end use due to high efficiency, low cost, and environmental stress by combining biological procedures. This paper reviews polymer nanocomposite membranes for water treatment.

Keywords Polymer nanocomposite · Membranes · Water treatment

1 Introduction

Every water molecule has two free electrons and two hydrogens to form four hydrogen bonds with other molecules and that exhibit the characteristics of water in solid and liquid state [1]. Water characteristic such as density [2] also depends upon temperature and pressure due to the presence of weak hydrogen bonds [3]. Hydrologic cycle is responsible for continuous circulation of water between earth's atmosphere, hydrosphere, and terrosphere by precipitation and evaporation, and fresh water is only three percent, thus playing a vital role in biological life [4]. With rising population and industrial development, the requirement of water has increased substantially. Therefore, many water treatment plants came into existence to purify contaminated water by applying biological and chemical methods due to the presence of metal and organic material, mainly classified as industrial, domestic, human, livestock exertion,

M. Sharma (✉)

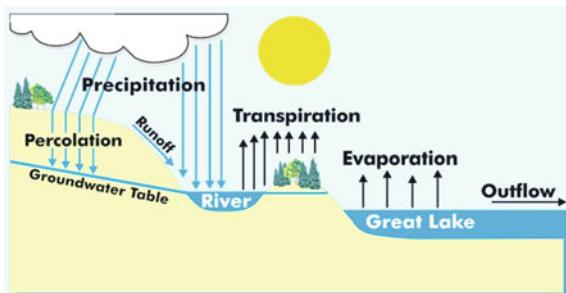
Department of Electrical Engineering, University of Petroleum and Energy Studies, Dehradun, India

e-mail: madhusharma@ddn.upes.ac.in

A. Sharma

Center for Energy Studies, Indian Institute of Technology Delhi, Hauz Khas, New Delhi, India
e-mail: anupamsharma.gec@gmail.com

Fig. 1 Hydrologic cycle [7]



etc. [5]. The inimical consequence on the bionomical arrangement by inclusion of wastewater should be avoided for the welfare of the society [6]. To preserve natural resources, water treatment is mandatory (Fig. 1).

2 Water Treatment Methodology

Membranes play a crucial role in wastewater recycling and advanced water treatment [8] and are widely used for desalting seawater [9], holding back seawater intrusion [10], treating brackish groundwater [11], water softening [12], wastewater recovery [13], removing color [14], odor [15], and other organic contaminants [16]. Membrane characteristic includes surface charge and roughness, pore size, water affinity, and material types [17]. Membranes with high and low pressure usually separate dissolved solids and suspended particulate matter, respectively, from water [18]. As compared to the conventional treatment plant, systems with membranes provide a high level of protection with low cost and reduce chemical use with less sludge disposal issues with possible remote operation [19]. Hollow fiber membranes that can be classified as microfiltration or ultrafiltration can be constructed from a range of materials that include polyvinylidene fluoride [20], polysulfone [21], polyacrylonitrile [22], polypropylene [23], polyvinyl chloride [24], polyethersulfone [25], etc. Surface water in either solid (glacier, snow cover) or liquid (lakes, rivers, bogs, temporary streams, reservoirs) form plays a crucial role in proper functioning of the ecosystem [26] and can be treated by membrane technology that includes microfiltration, bioreactors, etc. [27] (Figs. 2, 3, 4, 5, 6 and 7).

Fig. 2 Structure of polyvinylidene fluoride [28]

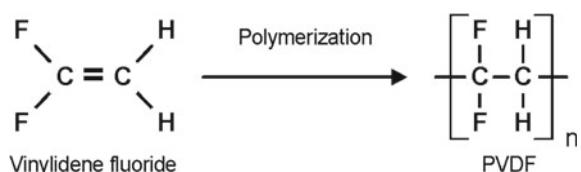


Fig. 3 Structure of polysulfone [29]

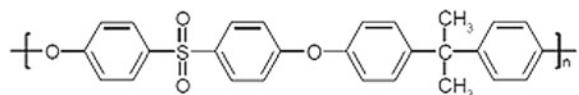


Fig. 4 Structure of polyacrylonitrile [30]

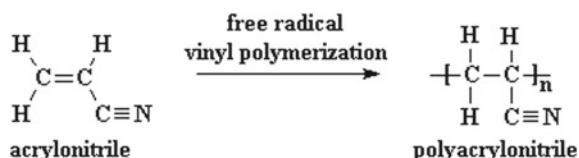


Fig. 5 Structure of polypropylene [31]

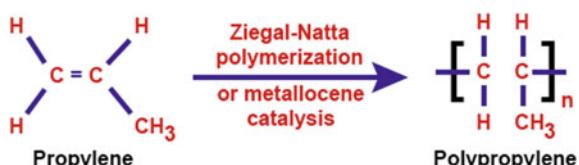


Fig. 6 Structure of polyvinyl chloride [32]

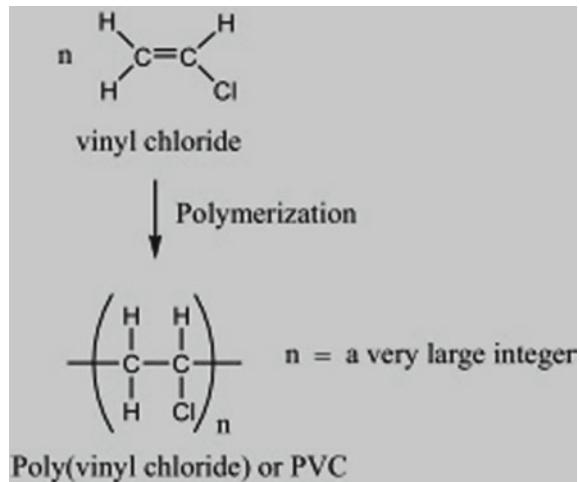
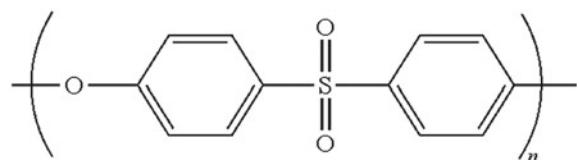


Fig. 7 Structure of polyethersulfone [33]



Fouling and concentration polarization are the factors that restrain the development of membrane technology [34]. Microorganism attached to the membrane surface leads to biofouling [35], which also depends on the quality of feed water and operating conditions. Mitigation of organic [36], inorganic [37], and colloidal [38] fouling is mandatory to increase the life of membrane [39] (Figs. 8, 9, 10, 11 and 12).

For a molecule separation, composite membrane materials, generally made with synthetic polymer and glass filler, are being developed to enhance membrane matrix molecular flux [45], which have better properties to those of components in pure form. At very low concentration, nanoscale fillers can improve the functional performance of immense polymeric material [46].

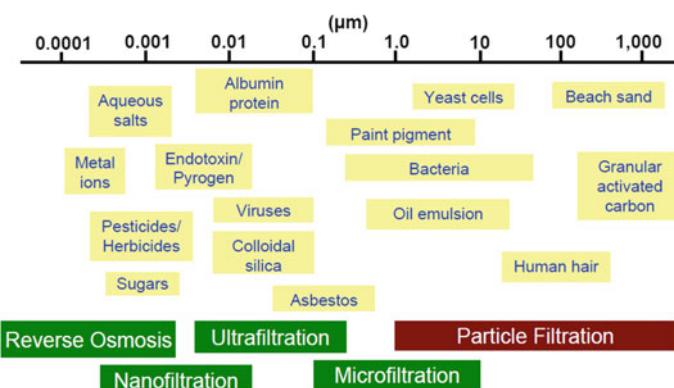


Fig. 8 Membrane characteristic [40]

Fig. 9 Characteristic of microfiltration membrane [41]

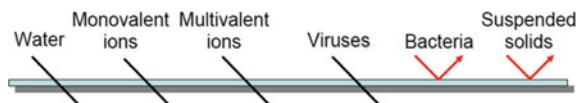


Fig. 10 Characteristic of ultrafiltration membrane [42]

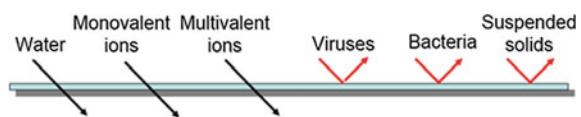


Fig. 11 Characteristic of nanofiltration membrane [43]

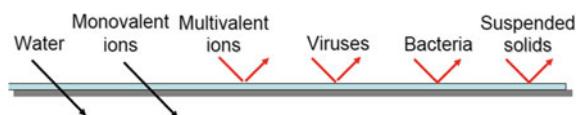
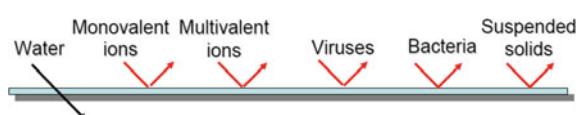


Fig. 12 Characteristic of reverse osmosis membrane [44]



3 Polymer Nanocomposites

Due to excellent separation properties [47] and economical operation [48], polymeric membranes are widely used with a wide range of flexibilities [49], whereas they have limited thermal, chemical, and mechanical resistance [50] and poor antifouling ability [51]. Composite polymeric membranes can be made to have very high strength [52] and permeability [53]. Preparation of polymeric nanocomposite membranes includes mixing of fillers in nanoscale and polymer matrix [54], widely used for gas separation [55], water and wastewater treatment [56], desalination [57] fuel cell [58] applications, etc. (Fig. 13).

Distribution and dispersion [60] of nanofiller determine the properties of polymer nanocomposite membranes [61]. Dispersion of nanoparticle can enhance physical [62], processing [63], and thermochemical [64] characteristics. Nanocomposite membranes incorporate nanoparticles such as titanium dioxide [65], silicon dioxide [66], aluminum oxide [67], silicon [68], silver [69], zinc oxide [70], zirconium dioxide [71], magnesium hydroxide [72], calcium carbonate [73], titanium silicon oxide [74], graphene oxide [75], carbon nanotubes [76], etc., and nanoparticle composites include $\text{SiO}_2\text{-TiO}_2$ [77], GO-SiO_2 [78], Ag-SiO_2 [79], GO-TiO_2 [80], etc. (Figs. 14, 15, 16 and 17).

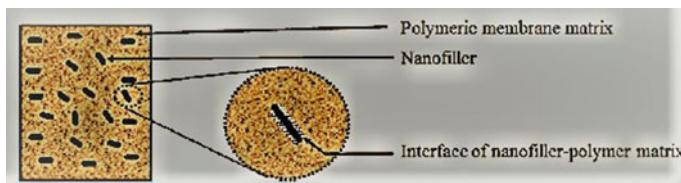


Fig. 13 Nanofillers in polymeric nanocomposite membrane matrix [59]

Fig. 14 $\text{SiO}_2\text{-TiO}_2$ nanocomposite structure [81]

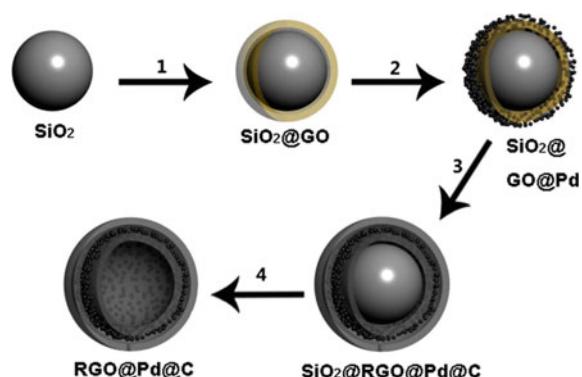


Fig. 15 GO-SiO₂ nanocomposite structure [82]

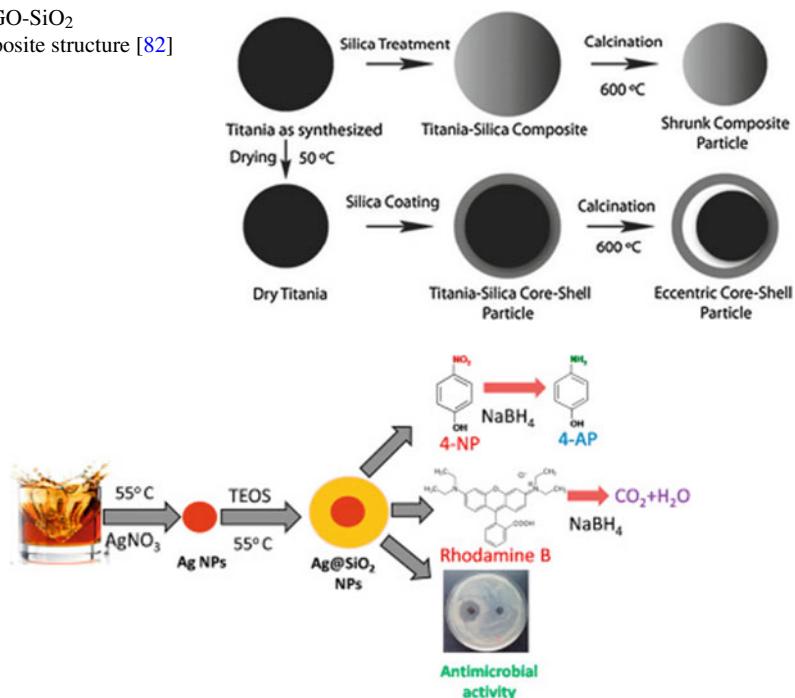


Fig. 16 Ag-SiO₂ nanocomposite structure [83]

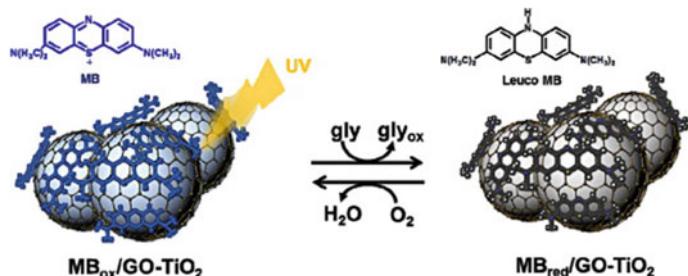


Fig. 17 GO-TiO₂ nanocomposite structure [84]

4 Conclusion

Clearly, nanomaterials can make future water treatment plants more effective, efficient, and cheaper to operate as well as making them quite easy to build in evolving countries where the supply of pure water is restricted.

References

1. Eisenberg, D., Kauzmann, W.: The structure and properties of water. Oxford University Press, London (1969)
2. Marion, C.M., Jakubowski, S.D.: Cold Regions Sci. Technol. **38**, 211–218 (2004)
3. Mishchuk, N.A., Goncharuk, V.V.: On the nature of physical properties of water. J. Water Chem. Technol. **39**(3), 227–240 (2017)
4. Michel, R.L.: Tritium in the hydrologic cycle. Springer, Berlin (2005) ISBN 10 1-4020-3010-X
5. Lee, L.K., Kim, J.H., Park, J., Kim, J.: Water quality at water treatment plants classified by type. Toxicol. Environ. Health Sci. **8**(5), 296–301 (2016)
6. Mullins, T.: The chemistry of water pollution, pp. 331–400. Springer, Boston (1977)
7. www.watershedcouncil.org. Accessed 2 Sept
8. Noeaghae, T., Kim, J.O., Chae, S.R.: Recent advances in Nano-Hybrid membranes for advanced water treatment. Curr. Org. Chem. **18**(18) (2014)
9. Boubakri, A., Hafiane, A., Bouguecha, S.A.T.: Direct Contact Membrane distillation: Capability to desalt raw water. Arab. J. Chem. **10**(supplement), 2 (2017)
10. Wintgens, T., Melin, T., Schafer, A., Khan, S., Muston, M., Bixio, D., Thoeye, C.: The role of membranes processes in municipal wastewater reclamation and reuse. Desalination **178** (2005)
11. Le, N.L., Nunes, S.P.: Materials and membrane technologies for water and energy sustainability. Sustain. Mater. Technol. **7**, 1–28 (2016)
12. Bodzek, M., Kotter, S., Wesolowska, K.: Application of membrane techniques in water softening process. Desalination **145** (2002)
13. Bolto, B., Hoang, M., Xie, Z.: A review of water recovery by vapour permeation through membranes. Water Resour. (Elsevier) (2012)
14. Cristiane, C.N., Petrus, J.C.C., Riella, H.G.: Color and COD retention by nano-filtration membranes. Desalination **172** (2005)
15. Hai, F.I., Yamamoto, K.: Membrane biological reactors. In: Wilderer, P. (ed.) Treatise on Water Science, pp. 571–613. Elsevier, Amsterdam (2011)
16. Nghiem, L.D., Fugioka, T.: Removal of emerging contaminants for water reuse by membrane technology. Emerging Membrane Technology for Sustainable Water Treatment, pp. 217–247. Elsevier, Amsterdam (2016)
17. Oliver Tania, Rania Ahmed, Joo Hwa: Membrane bioreactor technology for wastewater treatment and reclamation: membrane fouling. MDPI J. Memb. (2016)
18. Lorch, W.: Handbook of water purification. McGraw Hill, London (1981)
19. Lloyd: J. Membr. Sci. **52**(3), 239–261
20. Huang, Wang, Yan, Chen, Guo, Lang: Versatile polyvinylidene fluoride hybrid ultrafiltration membranes with superior antifouling, antibacterial and self-cleaning properties for water treatment. Elsevier, Amsterdam (2017)
21. Jeonghyun, Jungsik, Kim, J.M., Sang: Water treatment by polysulfone membrane modified with tetrahydrofuran and water pressure. Macromol. Res. **24**(11), 1020–1023 (2016)
22. Pasaoglu, Guclu: Polyethersulfone/polyacrylonitrile blended ultrafiltration membranes: preparation, morphology and filtration properties. Water Sci. Technol. (2016)
23. Sarah, Arshad, Nadeem: Fabrication and characterization of microfiltration blended membranes. J. Desalination Water Treat. **52**(10–12) (2014)
24. Behboudi, A., Zafarzadeh, Y., Yegani, R.: Polyvinyl chloride/Polycarbonate blend ultrafiltration membranes for water treatment. J. Membr. Sci. **534**, 18–24 (2017)
25. Celik, Park, Choi: Carbon Nanotube Blended Polyethersulfone Membranes for Fouling Control in Water Treatment. Elsevier, Amsterdam (2010)
26. Khublaryan, M.G.: Surface water: rivers, streams, lakes and wetlands, vol. 1, UNESCO, Encyclopedia of Life support systems
27. Lu-Li, Chettiyappan: Membranes technology for surface water treatment: advancement from microfiltration to membrane bioreactor. Review in Environmental Science and Bio/technology, pp. 1–24, (2017)

28. www.davis.com. Accessed 2 Sept
29. www.polymerdatabase.com. Accessed 2 Sep
30. www.pslc.ws. Accessed 2 Sep
31. www.alamy.com. Accessed 2 Sep
32. Emad & Ali, "Photo stabilization of poly vinyl chloride-still on the run", Journal of Tiabah university of Science, Vol. 9, Issue 4, Oct 2015
33. www.scielo.br. Accessed 2 Sep
34. Randon, Blanc, Paterson: Modification of ceramic membrane surface using phosphoric acid and alkyl phosphoric acids and its effect on ultrafiltration of BSA protein. *J. Membr. Sci.* (1995)
35. Thang, Felicity, Linhua: Biofouling of water treatment membranes: a review of the underlying causes, monitoring techniques and control measures. *Membranes (Basel)* **2** (2012)
36. Sunjin, Noeon, Sungyun, Jaeweon: Membrane characterization for mitigation of organic fouling during desalination and wastewater reclamation. *Desalination* **238**(1–3), 70–77 (2009)
37. Han, Zhou, Wan, Yang, Chung: Investigations of inorganic and organic fouling behaviors, antifouling and cleaning strategies for pressure retarded osmosis membranes using seawater desalination brine and wastewater. *Water Res.* **103**, 264–275 (2016)
38. Hong, Jin, Seung: A study on fouling mitigation using pulsing electric fields in electro dialysis of lactate containing BSA. *Korean J. Chem. Eng.* **19**(5), 880–887 (2002)
39. Wenshan, Huu-Hao, Jianxin: A mini-review on membrane fouling. *Bioresour. Technol.* **122**, 27–34 (2012)
40. Boussu, K., Vandecasteele, C., van der: Relation between membrane characteristic and performance in nanofiltration. *J. Membr. Sci.* **310**(1–2) (2008)
41. Reid, O.W.: Microfiltration membranes: characteristics and manufacturing. *Adv. Biochem. Eng.* (2006)
42. Heru, Mathias: Influence of ultrafiltration membrane characteristic on adsorptive fouling with dextrans. *J. Membr. Sci.* **266**(1–2), 132–142 (2005)
43. Timmer, J.M.K.: Properties of Nanofiltration Membranes-model Development and Industrial Application. *Technische Universiteit Eindhoven*, Eindhoven (2001)
44. Tolba, A.M., Mohamed, R.A.: Performance and Characterization of reversa osmosis membranes. In: Fourth International Water Technology Conference IWTC 99
45. Kubaczka, A.: Prediction of maxwell-stefan diffusion coefficients in polymer -multicomponent fluid systems. *J. Membr. Sci.* **470**, 389–398
46. Mondal, S., Hu, J.L.: Microstructure and water vapor transport properties of functionalized carbon nanotube reinforced segmented polyurethane composite membrane. *Polym. Eng. Sci.* **48**, 1718–1724 (2008)
47. Syrtsova, D.A., Kharitonov, A.P., Teplya, V.V., Koops, G.H.: Improving gas seperation properties of polymeric membranes based on glassy polymers by gas phase fluorination. *Desalination* **163**(1–3) (2004)
48. David, Zachary, Ruilan, Lloyd, James, Donald: Energy-efficient polymeric gas separation membranes. *Polymer* **54**(18) (2013)
49. Sangil, Francesco, Jerry: Fabrication of flexible, aligned carbon nanotube/polymer composite membranes by in-situ polymerization. *J. Membr. Sci.* **460**, 91–98 (2014)
50. Ibelnaweed, Javaid, S.M.: Thermal and mechanical properties of fuel cell polymeric membranes: strucure -property relationship. *Polymer Membranes for Fuel Cells*, pp. 235–252 (2009)
51. Lin, Mavis: Improved antifouling properties of polymer membranes using a layer-to-layer mediated method. *J. Mater. Chem.* (41) (2013)
52. Jamuna, K., Sabri, S.N.: Strength improvement of a functionally graded and layered composite membranes for guided bone regeneration in orthopedics. *Mater. Today Proc.* (2016)
53. Yuanyuan, Vicki: Enhancing membranes permeability for Co₂ capture through blending commodity polymers with selected PEO and PEO-PDMS copolymers and composite hollow fibers. *Energy Procedia* (2014)
54. Reddy, Mohan, Shah, Ghosh: Surface modification of ultrafiltration membranes by preabsorption of a negatively charged polymer. *J. Membr. Sci.* (2014)

55. Hailin, Maciej, Brian, Youqing: Polymer-inorganic nanocomposite membranes for water treatment. *Sep. Purif. Technol.* **55**, 281–291 (2007)
56. Yin, J., Ding, B.: Polymer matrix nano-composite membranes for water treatment. *J. Membr. Sci.* **479**, 256–275 (2015)
57. Jainesh and Z.V.P Murthy, “Nano composite membranes”, *Desalination and water treatment*, Vol. 57, Issue 55, 2016
58. Kim, D.J., Nam, S.Y.: A review of polymer-nanocomposite electrolyte membranes for fuel cell applications. *J. Ind. Eng. Chem.* **21**, 36–52 (2015)
59. Mondal, S.: Polymer nano composite membranes. *J. Membr. Sci. Technol.* (2014)
60. Sunhyung, Kyu, Bernd, Kyung, Christian: Structural development of nanoparticle dispersion during drying in polymer nanocomposite films. *Macromol. Am. Chem. Soc.* (2016)
61. Yin, J.: Fabrication and modification of nanocomposite membranes for enhanced water purification. University of Missouri (2014)
62. Paul, D.R., Robeson, L.M.: Polymer nanotechnology: nanocomposites. *49*(15), 3187–3204 (2008)
63. Kerstin, Elodie, Marcos, Markus: Review on the processing and properties of polymer nanocomposites and nanocoatings and their applications in the packing, automotive and solar energy fields. *J. Nanomater.* (2017)
64. Ben, Dan, Jing, Hong: Rubbery polymer-inorganic nanocomposite membranes: free volume characteristic on separation property. *Ind. Eng. Chem. Res.* **49**, 12444–12451 (2010)
65. Devrim, Y.: Preparation and testing of Nafion/titanium dioxide nanocomposite membranes electrode assembly by ultrasonic coating technique. *J. Appl. Polym. Sci.* (2014)
66. Suk-Yong, Sien-Ho: Sulphonated polySEPS/hydrophilic-SiO₂ composite membranes for polymer electrolyte membranes. *J. Ind. Eng. Chem.* **23**, 285–289 (2015)
67. Yasaman, Mohammad, Maryam: Fabrication of AL2O3/PSF nanocomposite membranes: efficiency comparison of coating and blending methods in modification of filtration performance. *J. Desalination Water Treat.* **51**(34–36) (2013)
68. Vajiheh, May-Britt: Development of nanocomposite membranes containing modified Si nanoparticles in PEBAK-2533 as a black copolymer and 6FDA-durene Diamine as a glassy polymer. *ACS Appl. Mater. Interfaces.*, 15643–15652 (2014)
69. Jiang, Y., Liu, D., Minjung, Pratim, John: In situ photocatalytic synthesis of Ag nanoparticles by crumpled graphene oxide composite membranes for filtration and disinfection applications. *Am. Chem. Soc.* (2016)
70. Robin, Hruda, Nandakumar: Electrospun polycaprolactone/ZnO nanocomposite membranes as biomaterials with antibacterial and cell adhesion properties. *J. Polym. Res.* (2014)
71. Jingjing, Haining, Wei: Nafion-zirconia nanocomposite membranes formed via in situ sol-gel process. *Int. J. Hydrog. Energy* **35**(7), 2796–2801 (2010)
72. Huaqiang, Jiefu, Baojun: Mg(OH)₂ complex nanostructures with super hydrophobicity and flame retardant effects. *J. Phys. Chem.* (2010)
73. Sanaz, Samira, Nor, Wan, Mohamad, Susan: Enhancement of mechanical and thermal properties of polycaprolactone/chitosan blend by calcium carbonate nanoparticle. *Int. J. Mol. Sci.* **13**, 4508–4522 (2012)
74. Yilser, Serdar, Nurcan, Inci: Nafion/titanium silicon oxide nanocomposite membranes for PEM fuel cells. *Int. J. Energy Res.* (2012)
75. Swati, Prem, Hariom, Vaibhav, Prafulla: Preparation of graphene oxide nano-composite ion-exchange membranes for desalination applications. *R. Soc. Chem.* (2014)
76. Hosam, So-Ryong, Shihong, Mark: Synthesis and characterization of a carbon nanotube/polymer nanocomposite membrane for water treatment. *Desalination* **272**(1–3) (2011)
77. Zhao, L., Jiaguo, Bei: Preparation and characterization of SiO₂/TiO₂ composite microspheres with microporous SiO₂ core/mesoporous TiO₂ shell. *J. Solid State Chem.* **178**(6), 1818–1824 (2005)
78. Nanying, Minglu, Jing, Liqun, Ming: Simultaneously improved actuated performance and mechanical strength of silicone elastomer by reduced graphene oxide encapsulated silicon dioxide. *Int. J. Smart Nano Mater.* **6**(4) (2015)

79. Young, Don: Synthesis and characterization of Ag and Ag–SiO₂ nanoparticles. *Collides Surf. Physicochemical Eng. Asp.* **257–258**, 273–276 (2005)
80. Guodong, Chao: TiO₂ nanoparticles assembled on graphene oxide nano sheets with high photocatalytic activity for removal of pollutants. *Carbon* **49**(8), 2693–2701 (2011)
81. Harrison, Rui: Versatility of heterogeneous photo catalysis: synthetic methodologies epitomizing the role of silica support in TiO₂ based mixed oxides. *R. Soc. Chem.* (2012)
82. Zheye, Tai: Encapsulating Pd nanoparticles in double-shelled graphene @carbon hollow spheres for excellent chemical catalytic property. *Sci. Rep.* **4**, (2014)
83. Otari, S.V.: Facile one pot synthesis of core shell Ag@SiO₂ nanoparticles for catalytic and antimicrobial activity. *Mater. Lett.* **167**, 179–182 (2016)
84. Eun-Jin, Joon: Self-adhesive graphene oxide-wrapped TiO₂ nanoparticles for UV activated calorimetric oxygen detection. *Sen. Actuators Chem.* **213**, 322–328 (2015)

Analysis and Comparison of ECG Signal Quality Assessments Methods



Ramesh Babu Chukka and Ch. Sumanth Kumar

Abstract An ECG, SQA plays an important role in significantly improving the performance. However, in real cases, artifacts often corrupt. In this paper, several methods have been analyzed for ECG enhancement like fast Fourier transforms (FFT), empirical mode decomposition (EMD) and Hilbert vibration decomposition (HVD) based filtering. These methods are able to remove both high frequency and baseline wander noises. The method is validated through experiments on the PhysioNet/MITBIHA database records. Techniques are compared in terms of correlation, SNR and detection error rate. The simulation shows that the HVD-based filtering method provides a very good result for denoising and BW removal. In addition, we analyzed two types of SQA algorithms features and rules-based method and filtering and threshold-based signal quality assessment methods, which are intended to assess reliable heart rates and the good quality of ECG signal. The proposed methods are well suited for transmitting valid data for cloud diagnosis.

Keywords Electrocardiogram · ECG noises · Signal quality assessment

1 Introduction

Health is a primary need in our life. We are facing problems in getting proper health care because of limited technology and resources. All of ECG applications such as cardiovascular diseases diagnosis, arrhythmias recognition, sudden cardiac arrest prediction, highly demands the exact determination of ECG signal [1].

R. B. Chukka (✉)

Department of Electronics and Communication Engineering, Vignan's Institute of Engineering for Women, Visakhapatnam, India

e-mail: rameshbabuchukka@gmail.com

Ch. S. Kumar

Department of Electronics and Communication Engineering, GITAM Institute of Technology, GITAM Deemed to Be University, Visakhapatnam, India

Challenges faced in the field of long-term healthcare monitoring include portability in size, weight, and low cost [2]. Noise-free signals are required for analysis, therefore, highly demanded for SQA in reducing the false alarms. The corrupted information may lead to the wrong diagnosis.

In this paper, we aimed to remove of baseline wander and high-frequency noises based on FFT, EMD, and HVD-based filtering technique is introduced. In addition, quality assessment strategies have been employed to tackle false alarms problems.

2 Methods and Methodology

For the implementation, ECG records are taken from standard databases as shown in Fig. 1.

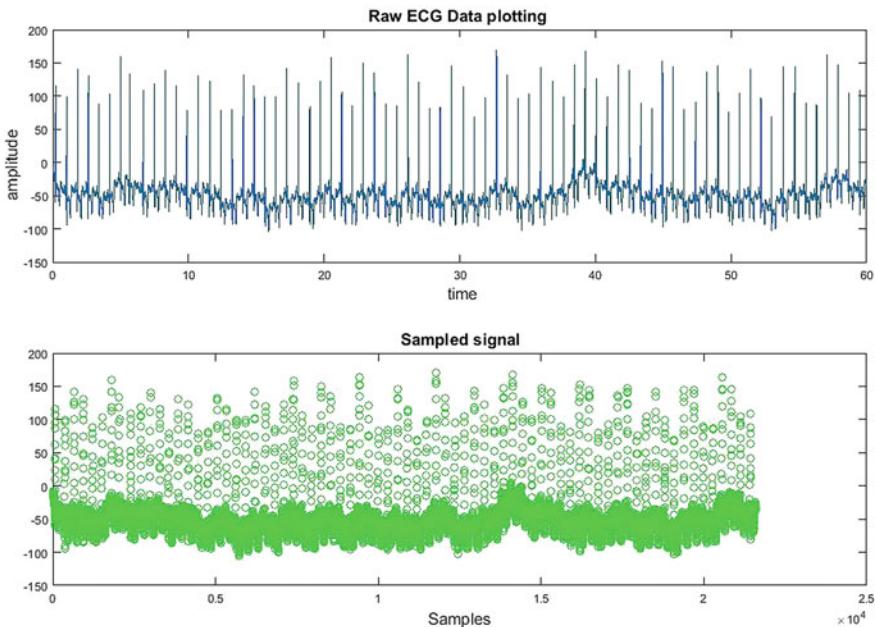


Fig. 1 Raw ECG data plot and sampled signal

2.1 ECG Denoising and Component Analysis

2.1.1 DWT-Based Filtering

Wavelet transform (WT) provides more information and can perform analysis of both frequency and time representation. An algorithm is used to compute the discrete WT (DWT) decompositions using LPFs and HPFs [3]. Figure 2 shows the BW removal using wavelet decomposition.

2.1.2 EMD-Based Filtering

An empirical mode decomposition (EMD) is introduced by Huang is used to analyze both time–frequency analysis.

Given an input signal $x(t)$, the effective algorithm of EMD can be summarized as follows:

1. Identify all IMFs of $x(t)$;
2. Generate the upper and lower envelope by connecting the maxima and minima;
3. Compute the local mean;
4. Extract the details;
5. Iterate on the residual.

The EMD method expresses the signal $x(t)$ as the sum of a finite number of IMFs and a final residual is obtained at the end of process.

The output of the EMD Technique is shown in Fig. 3.

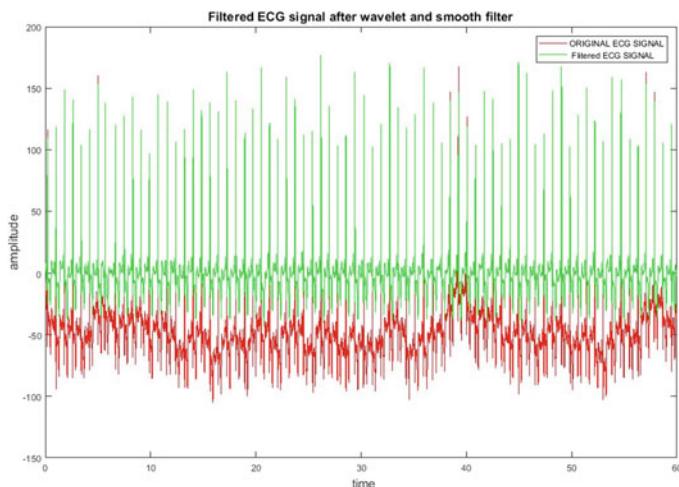


Fig. 2 DWT-filtered signal

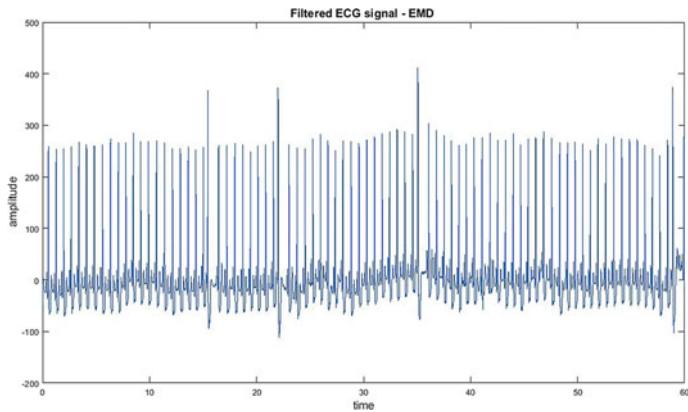


Fig. 3 EMD-filtered signal

2.1.3 HVD-Based Filtering

For better understanding, let we go with some mathematical issues.

(a) Estimation of the instantaneous frequency:

Let us consider a signal $x(t)$ can be expressed as

$$x(t) = \sum_l a_l(t) \cos\left(\int \omega_l(t) dt\right)$$

where

$a_l(t)$ is the instantaneous amplitude and

$\omega_l(t)$ is the instantaneous frequency of l component.

(b) The Envelope Detection:

We choose a technique well-known synchronous detection to extracts the amplitude details with a known frequency.

(c) Subtraction of the largest component:

Subtract the largest component from the initial signal.

$$x_{l=1}(t) = x(t) - x_1(t)$$

The outputs of the HVD Technique are shown in the figures (Fig. 4).

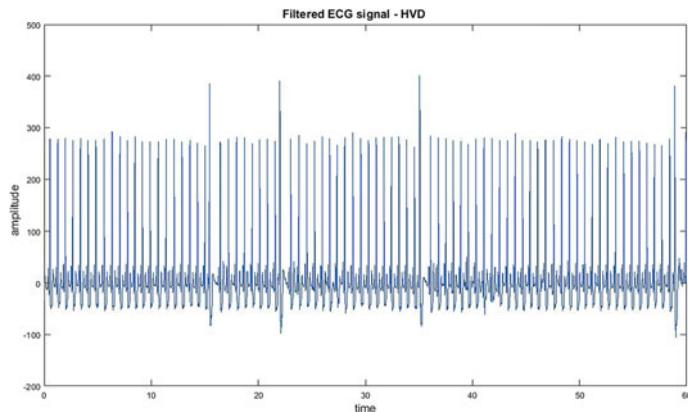


Fig. 4 HVD-filtered signal

2.1.4 Performance Metrics Comparisons

The three denoising techniques discussed above are applied on some of the recorded ECG signals available in the database, and their performances which are compared in terms of correlation criterion and SNR are tabulated.

PhysioNet database	DWT		EMD		HVD	
	Correlation (γ)	SNR	Correlation (γ)	SNR	Correlation (γ)	SNR
100	0.97	26.1	0.98637868	27.1229	0.99971302	45.9956
101	0.93	23.2	0.94212359	24.4681	0.99758573	38.5399
102	0.98	27.6	0.99739324	28.2829	0.99939312	38.7167
105	0.98	30.1	0.99110905	31.9607	0.99969329	44.2901
109	0.97	17.2	0.98532469	18.8084	0.99944116	36.879
111	0.97	26.4	0.98566024	27.1252	0.9991882	37.8521
112	0.94	20.1	0.95288089	21.8086	0.99483184	29.034
114	0.94	20.9	0.95746516	21.6429	0.99844693	27.316
115	0.97	27.6	0.98223207	28.4056	0.99895625	37.6982
116	0.86	25.2	0.87563436	26.8662	0.99259015	27.5232
118	0.94	23.6	0.95692183	24.966	0.99862837	29.7275
119	0.98	24.9	0.9912986	25.379	0.99839035	34.5063
121	0.66	17.1	0.67770319	18.1411	0.93164784	20.5073
233	0.98	21.5	0.99447502	22.515	0.99981577	39.4077
234	0.92	19.8	0.93065476	20.88	0.99969609	46.2976

From the above results, the overall performance of the HVD technique for removal of BW is significantly improved compared with the EMD technique [5] for the same BW signal. Thus, the proposed HVD technique is computationally efficient and better compared to EMD and DWT technique.

2.2 *Signal Quality Assessment's*

It is analyzed in two SQA methods (A) Features and rules based and (B) Filtering-based SQA method. This method consists of filtering stage, feature extraction, classification stage.

2.2.1 Features and Rules-Based Method

To determine the reliable heart rates by signal quality indices (SQIs). The signal quality was assessed by using features, rules and adaptive template matching [4].

2.2.2 Filtering and Threshold-Based Assessment Methods

Filters and Threshold-based Combined SQIs based on ECG quality estimator as shown in Fig. 5 is used for estimating the HR from the ECG signals [1].

2.2.3 Results and Discussion

To calculate the performance of quality assessments methods uses manual and automatic machine learning approaches to identify acceptable signals in terms of Sensitivity, Specificity, and Accuracy. Results show that features and rules-based method give 87.4% of Se%, 100% of specificity, and 90.2% of accuracy. Whereas, Filtering and threshold-based signal quality assessment methods had accuracy 94.2%, Sp of 100% and Se of 92.4%.

3 Conclusion

This paper presents analyzing of denoising filtering techniques. Overall performance is calculated in terms of correlation and SNR. In addition, this paper also presents ECG SQAs methods for acceptable ECG signal. Moreover, performance evaluations were only carried out for accurate identifications. For future extension, these methods are useful for real-time evaluation for telemetry healthcare monitoring applications.

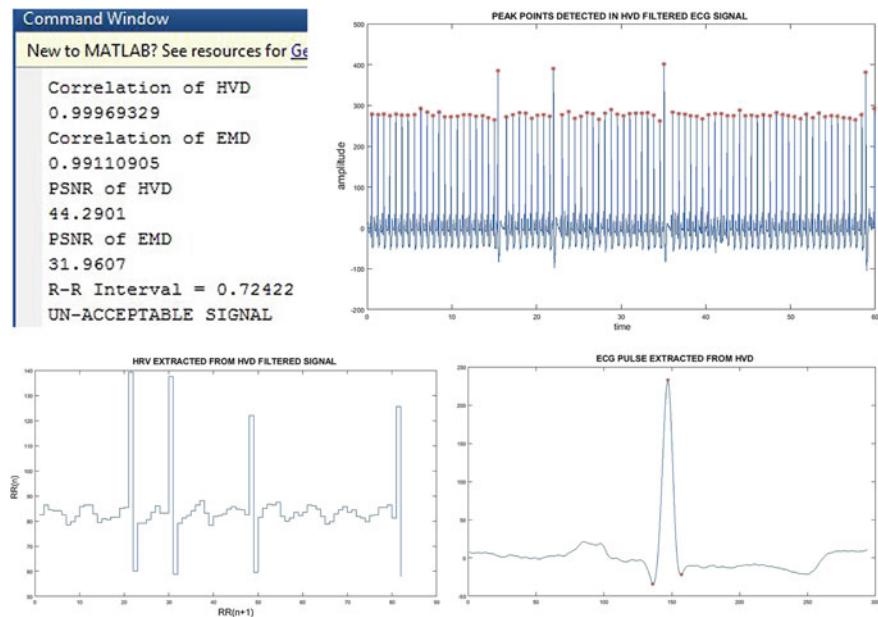


Fig. 5 Computation performance of SQA, peak detector, HRV, ECG pulse

References

1. Satija, U., Ramkumar, B., Manikandan, M.S.: Real-time signal quality-aware ecg telemetry system for iot-based healthcare monitoring. *IEEE Internet Things J.* **4**, 815–823 (2017)
2. Chukka, R.B., Kumar, C.S.: A novel architecture for IoT enabled ECG signal quality assessment using HVD based on PSoC System: a survey. *J. Adv. Res. Dyn. Control Sys.* **10**, 333–339 (2018)
3. Chukka, R.B., Kumar, C.S.: A novel architecture for the realisation of iot-enabled ECG signal quality assessment using wavelet decomposition for baseline wander removal. *Def. S&T Tech. Bull.* **11** (2018)
4. Christina, O., Timothy, B., Peter, C., David, C., David, V., Lionel, T.: Signal quality indices for the electrocardiogram and photoplethysmogram: derivation and applications to wireless monitoring. *IEEE J. Biomed. Health Inform.* **3**, 832–838 (2015)
5. Chukka, R.B., Kumar, C.S.: A novel architecture for the realisation of IoT-enabled ECG signal quality aware using empirical mode decomposition for telemetry healthcare system. *IJET(UAE).* Accepted for Oct 2018 Issue

Mobile Sensors Deployment Methods: A Review



P. Thrilochan Sharma, Sushabhan Choudhury and Vinay Chowdary

Abstract Usage of wireless sensor networks has become inevitable in the field of international border monitoring and zone surveillance. This application of sensors requires a barrier of sensors along the border to be monitored. The more the number of barriers, the better the monitoring system. Usage of a number of sensor nodes (SNs) in a specific configuration leads to the formation of a specific number of barriers. The challenge here is to develop a deployment algorithm, which provides the maximum number of barriers with the minimum number of SN relocation. Wireless Sensor Networks are used in hazardous environments and are preferred to be aerially deployed. In this paper, various methods of Mobile Sensor deployment methods are discussed along with their pros and cons. A comparative study is done based on many factors such as relocation distance, obstacle avoidance, scalability, etc.

1 Introduction

K-barrier coverage is one of the most important problems in Wireless Sensor Networks (WSNs). Widely used applications of k-barrier coverage are Intruder detection and border surveillance. All intruders' cross-paths through a region will be caught by at least k distinct sensors. K-different chain of overlapping sensors must be formed in WSNs to achieve effective k-barrier coverage. Usage of mobile sensors includes initial cost as well as maintenance cost by taking energy consumption for mobility into consideration. However, in general, barrier coverage cannot be satisfied after the initial deployment of sensors, hence coverage related problems of WSNs have gathered extensive attention.

P. Thrilochan Sharma
Automation and Robotics, University of Petroleum and Energy Studies, Dehradun, India

S. Choudhury · V. Chowdary (✉)
Electrical and Electronics Department, University of Petroleum and Energy Studies, Dehradun,
India
e-mail: vchowdary@ddn.upes.ac.in

Tri and Chakchai formulated a distributed deployment algorithm for improved barrier coverage with mobile sensors by relocating sensor nodes (SNs) [1]. He and Shi devised an efficient algorithm to construct k-barrier coverage with minimum cost and proposed a distributed algorithm to find the maximum barrier coverage in WSNs [2]. A centralized wakeup/sleep scheme was proposed to form barrier coverage. A distributed algorithm to construct multiple disjoint barrier path for strong barrier coverage was studied [3]. Kumar et al. classified barrier coverage as strong barrier coverage and weak barrier coverage [4]. Chang et al. concentrated to prolong barrier lifetime by achieving k-barrier coverage in terms of energy consumption of participating mobile sensors [5]. Based on distributed learning technique, Mstafaei proposed an efficient scheduling method to obtain near-optimal solution to the barrier coverage problem [6].

According to the accessibility of the area under surveillance, two types of sensor deployment techniques are defined, Random deployment or Deterministic deployment. When the monitored area is a human-controlled environment, deterministic sensor deployment shall be used. But in most cases, the area under surveillance is an inaccessible area and hence only random deployment of sensors shall be used.

In random deployment, SNs that are redundant, not belonging to any barrier, are in a considerable number. Hence by decreasing this number, and hence, saving a considerable amount of energy and thereby saving cost is the main challenge. All the methods discussed approaches utilize mobile sensors and apply centralized deployment algorithm to improve k-barrier coverage.

Many of the deployment approaches use clustering technique. In the clustering technique, the exchange of information is only limited to the cluster heads which reduces the information exchange between individual sensors resulting in loss of information. We have tried our best to study as many models as possible which try to nullify this kind of information loss and also saving energy, indirectly saving cost.

In particular, any deployment algorithm has to consider the following problems

- i. Effective k-barrier coverage with a minimum number of sensor relocations: Barrier coverage cannot be satisfactory due to some of the reasons like the formation of barrier gaps. SNs are relocated to reduces these gaps. But these relocations consume most of the energy in the network. Hence, we study various models trying to reduce the number of relocations.
- ii. To achieve k-barrier coverage reducing cost: This problem mainly involves achieving efficient coverage with minimum cost. Less movement of SNs means less energy consumed for movement.

The main objective of our research is

- We discussed various types of algorithms which use different approaches in deploying mobile sensors.
- To compare the characteristics and results of the suitable methods and present it in a comparative form.

2 Types of Models

As there is a lot of research going on about sensor deployment schemes, a various number of deployment techniques have been proposed. Each method has its own advantages as well as disadvantages.

1. Virtual Force Method [7]: This method is based on the clustering technique. In a cluster, an MSN is chosen as Cluster Head (CH) based on their random physical location. This CH manages other MSNs in the cluster. Each of the SNs exerts a force either attractive or repulsive, hence this method is called Virtual Force Method.

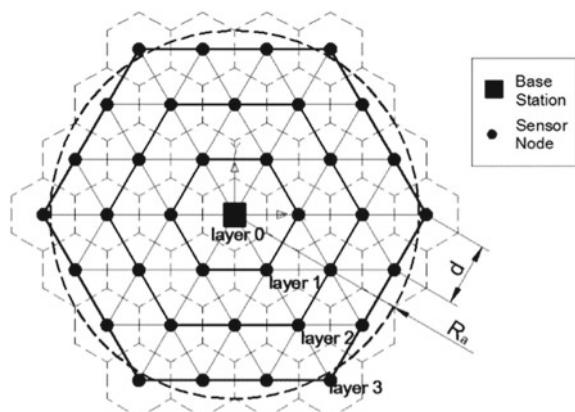
When an MSN is closer to another MSN, lesser than the defined value, then one SN exerts a repulsive force on the other. In the case of two MSNs getting farther, more than the predefined value, an attractive force starts acting in between them. This makes sure that the whole area is covered with maximum coverage of sensors.

But, a high amount of energy and time is consumed for the action of these forces. These forces are not physically applied. Cluster heads apply them logically deciding the final destination of each MSN.

2. Connectivity Preserved Virtual Force [8]: MSNs are randomly spread in a region. CPVF helps in ensuring the connectivity of all the randomly spread MSN. In this method, a Base Station (BS) will be allotted. All the SNs will consider this BS as a reference. The moment of all these MSNs will be decided by the path towards the BS from the SN.

In this method, obstacle avoidance is performed. In this method, SNs follow BUG2 algorithm (Right-Hand Rule) for handling obstacles. In this algorithm, the SN moves around the obstacle such a way that the right-hand touches the obstacle until a reference line has been detected. This method reduces moments making sure the network connectivity, but this method lacks stability (Fig. 1).

Fig. 1 Arrangement of base station and sensor nodes in CPVF



3. FLOOR Method [8]: This method is developed from CPVF. Similar to CPVF, BUG2 algorithm is used in the FLOOR for obstacle avoidance.

Initially, Floor lines are formed which divide the region into several regions. In this when an SN connects to Base station, it will be given IDs of its ancestors. This helps to determine which expansion of the coverage can be done by moving an MSN without disturbing the whole network.

Floor Head is elected for a floor which has the least value of x-coordinate. All the information of the MSNs present on that floor is handled by the Floor Head. After all the MSNs are located, the possible movements of MSNs to increase the effectiveness of coverage are determined. This method minimizes MSN movements but This method fails to achieve 100% coverage since some uncovered patches are formed by SNs of floor lines adjacent to it.

4. Push–Pull-Based Distributed Deployment Method (PPDD) [9]: This is a homogenous deployment method. In this method, Target region is divided into several hexagonal segments with MSN being the center of each segment, called tiling. At random time intervals, Tiling is initiated with MSNs present at arbitrary positions. Adjacent MSNs are also placed at the center of neighboring tiles taking the previous MSN as the initial point. This is called Snapping activity.

When an MSN is inside the hexagonal tile of another MSN, it is called slave. These MSNs are pushed from a high-density region to low-density region. This is called Push activity. After Push activity, some open spaces exist. These regions are covered by Pull activity initiated by snapped MSN triggering a message demanding a slave. Tile Merging is an activity of bringing together all the sub-regions in a target region. This activity is performed simultaneously with other activities.

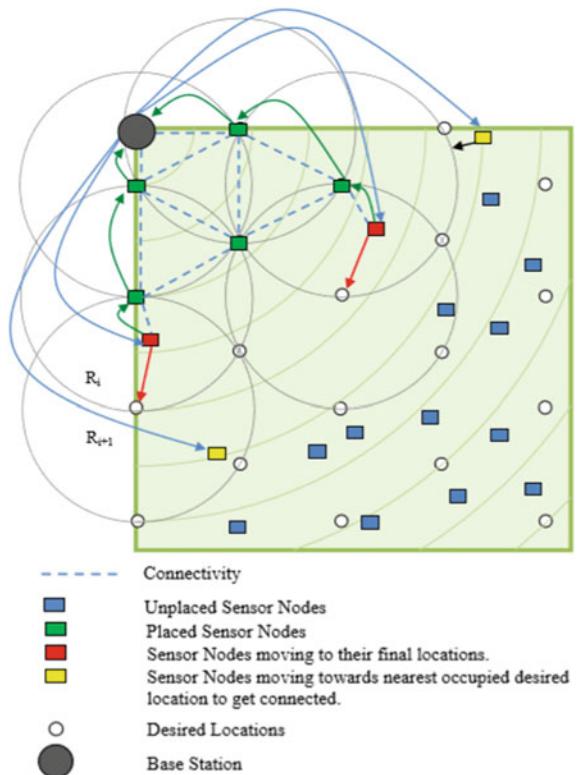
5. Distributed Deployment Scheme (DDS) [10]: This is a homogenous distribution method within a target region. In this method, BS is employed which logically divides the entire target region into various concentric regions with BS as their centers. BS computes all the points to place the MSNs for effective coverage.

After the location of all the points, BS calls MSNs to spread layer after layer starting from the inner circle. If there is no MSN in a layer's vicinity, BS holds for some time and invites MSNs from the next layer to connect.

This is an iterative process until all desired locations are allotted. This method facilitates 100% coverage with minimizing MSN movement and with minimum overlap, it facilitates multipath connectivity but due to layer-wise movement, this method takes a large amount of time (Fig. 2).

6. Scalable Energy-Efficient Deployment Scheme (SEEDS) [11]: This method is similar to DDS. It is a scalable technique which was designed to maximize the coverage with a minimum number of SNs and minimum relocation. Both DDS and SEEDS use the same technique for determining placements of SNs. They divide the target region into several hexagon tiles. Even the same communication and sensing range relation is followed.

Fig. 2 Movement of sensor nodes in DDS [10]



It differs from DDS due to its scalability and distribution algorithm used for movement of MSNs to Desired Locations (DLs). It also includes an obstacle handling algorithm. BS sends the total list of predetermined DLs to all MSNs which are scattered randomly. Then the movement of these MSNs toward each DL in the specified sorted list until all the DLs are occupied. After an MSN reaches a DL, it sends a stop message in order to stop another MSN approaching the DL. The disadvantage of the method is that as the target region increases, the size of DL list also increases and also the size of computation overhead and broadcast packet increases as they sort this list depending on their current position.

7. Distributed Self-Spreading Algorithm Method [12]: DSSA is a Sensor Node deployment method. In this, SNs are relocated repeatedly within a target region until the desired density is reached throughout the region. DSSA is proposed based on the density and space between the molecules in a compound.

Every MSN makes decisions independently regarding their movement both in terms of magnitude and direction. These decisions are made based on two factors. One is the neighboring MSNs exerted force and the other is the density difference

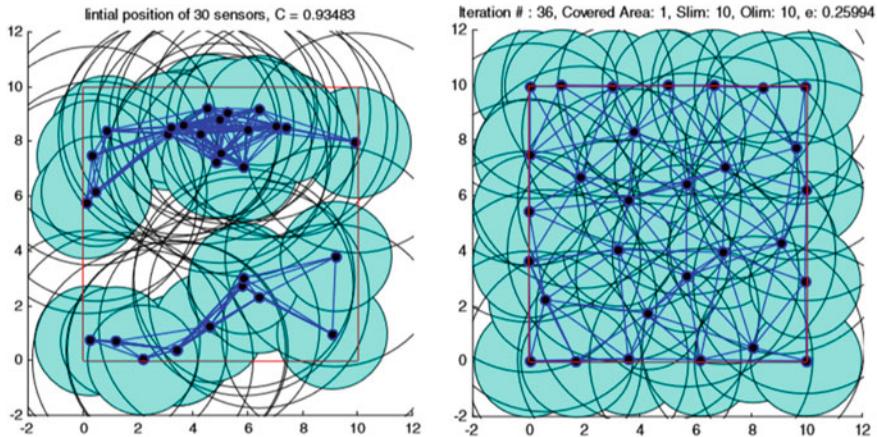


Fig. 3 Sensor nodes initial and final arrangement applying DSSA [12]

in a given area between local and reference values. DSSA minimized the MSN movements effectively but it does not guarantee BS connectivity (Fig. 3).

8. Centroid-Based Movement Assisted Deployment Method (CBMA) [13]: CBMA is proposed by the authors in [13] along with dual centroid based method. It proposes an algorithm for homogeneous deployment of SNs in a target region.

Both methods are iterative which relocate the MSNs to the centroid of Voronoi's polygon formed by them. Coverage enhancement after the relocation is taken as a criterion to check the benefit of relocation.

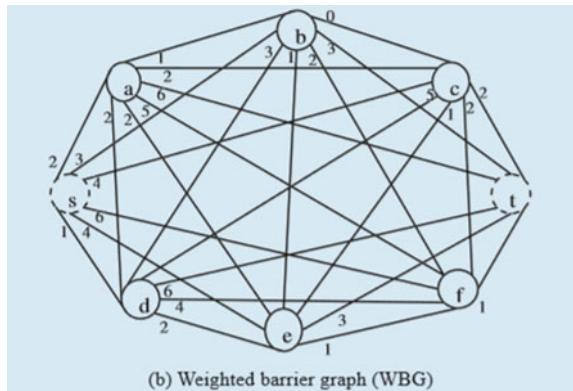
Centroid method is applied in several phases. At the starting position of every phase, Local Voronoi's polygon is formulated by each and every MSN with the help of the positional information of adjacent MSNs. Each MSN computes a new centroid if the coverage of its local polygon is incomplete.

Dual Centroid method is an expansion of the centroid method. Instead of a single centroid, using the polygon formed by adjacent MSNs, two centroids positions are located and relocation position is determined.

9. Glow-Worm Swarm Optimization Algorithm (GSO) [14]: Taking Inspiration from Glow-worm's behavior, this method was formulated. Glow-worms emit luciferin which makes them glow with intensity varying with distance. Other glow-worms move towards the neighboring glow-worm with higher intensity and form clusters. In GSO, converse to glow-worm technique, MSNs are uniformly distributed within the target region.

Position Coordinates of every MSN is well known and with the help of these coordinates, the distance from the neighboring MSN is calculated by itself. The magnitude of the distance is the factor in determining the luciferin intensity of the MSN. Based on this intensity, MSN calculates the possible movement in the direction of the neighboring MSNs and MSN with the highest probability is selected. MSN

Fig. 4 Weighted barrier graph [15]



starts moving towards it. This process is repeated until the target region is covered uniformly.

10. RSMN, RELAX-RSMN, UNRES-RSMN [15]: These methods are based on the network model of Weighted Barrier Graph (WBG). RSMN stands for Relocate Sensors with Minimum Number. As the name suggests, the main function of these models is reducing the number of relocating sensors as much as possible.

Weighted Barrier Graph (WBG) is a novel model introduced by Wang et al. [16]. WBG is described using triple $\langle V, E, W \rangle$. Here V represents the sensor node set. In V , there are two virtual nodes corresponding to the left and right boundaries namely, s and t . RSMN problem is the formation of k -sensor barriers with a minimum number of relocated sensors. The problem is formulated in Integer Linear Programming model (ILP) (Fig. 4).

Because of the integrality and complicated constraints, RSMN is computationally intractable. The RELAX-RSMN problem statement is formulated by solving the computationally intractable problem after relaxing the integrality constraints. UNRES-RSMN model is the same as the RSMN model except for the total barrier sensor constraint.

11. Vector-based Algorithm [17]: The approach of this method is push based. Three algorithms have been proposed by authors in [17] for uniform distribution of MSNs in a target region. Holes in the candidate region are detected using Voronoi's graph in all the three algorithms (VEC, VOR, and MINIMAX). Regions having sparse MSNs exert pulling force and regions with dense MSNs exert pushing force. This results in the expansion of coverage in the target region. In VEC, the forces between MSNs and between MSN and target region boundary is assumed to be similar to the repulsive force among two similarly charged particles.

Average distance (d_{avg}) is computed initially and considered as standard between MSNs. Repulsive force can be observed if the distance between MSNs is less than d_{avg} and if it is more than d_{avg} then the attractive force is exerted between MSNs.

Coverage after a proposed movement is computed beforehand to control the unnecessary movement of MSNs. Voronoi's polygon is used for computing coverage at any point in the sensing region of MSN. The point at 1/2 distance or ¾ distance of the total distance is taken as the target position, if the target region is too far.

12. Voronoi-Based Algorithm [17]: Voronoi approach (VOR) is a pull dependent approach. MSN tries to fill holes present in the voronoi's polygon by carrying out movement towards the farthest vertex in this method. Displacement restriction is used to control the movement of MSNs. A new Voronoi's polygon is drawn as MSN discovers new neighbors while its movement. By avoiding immediate backward movement and keeping a record of previous move oscillations are controlled.

13. Minimax [17]: Minimax algorithm is similar to VOR. In Minimax algorithm MSNs move towards the farthest vertex to cover the holes in the Voronoi's polygon. But it abides by a condition that while moving closer to a target location, closer vertices do not become furthest. And the distance between the farthest vertex and MSN is maintained to be minimum. Minimax point is the computed final point for relocating the MSN. All the three algorithms displace MSNs to more appropriate positions to enhance the coverage, thus consuming more energy.

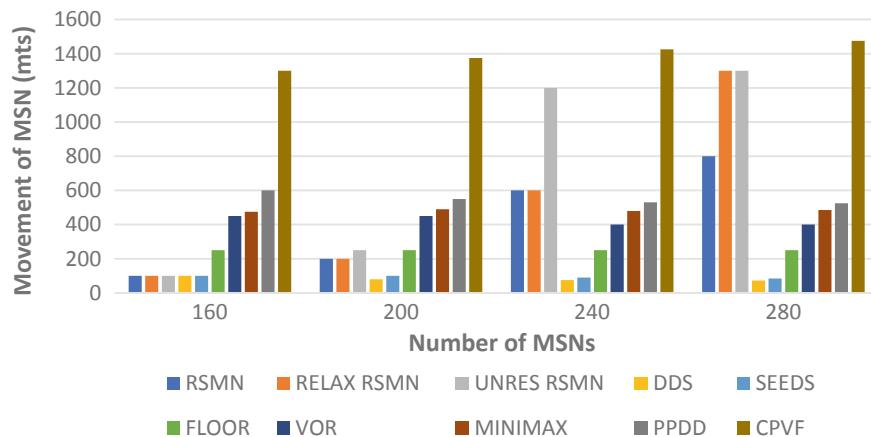
3 Analysis and Discussion

Major part of the energy consumed in any of the above-discussed models is taken for the movement of MSNs. Various researches are going on to minimize MSN movements. Comparative study of various methods is depicted in Fig. 5. Sophisticated nature of MSNs is the reason for them to have limited accessibility and mobility based on working terrain.

Some methods such as PPDD, DDS, FLOOR, SEEDS, CPVF enhance the network stability guaranteeing the connectivity of SNs with the BS, while some other methods focus mainly on uniform distribution of SNs within the target region.

Comparative analysis can be seen in Table 1. Scalability generalizes the model to any size of the target region making it independent of other variables. Hence, the scalability of a model is important. DDS and SEEDS aren't scalable while other methods are scalable. Some of the models like PFM, VFM, CPVF, PPDD, SEEDS, FLOOR include obstacle-handling algorithms in target region which is more realistic in nature while other models consider ideal target region (Plane surface without obstacles). Target regions are mostly inaccessible with uneven terrain and hazardous environments making ground deployment difficult. Moreover, most of the methods considered aerial dropping for the random scattering of the MSNs but did not consider the wind effect on the dropping.

When an SN has dropped aerially from a helicopter, the potential energy owned by the SN is converted into kinetic energy which will adversely affect the SN. Use of parachute or a spongy base to act as buffers against the shocks and vibrations are considered. Bio-inspired deployment models inspired by analyzing flight patterns of

**Fig. 5** Comparative study of schemes**Table 1** Comparative analysis of deployment models

Model	Sensor type	Connectivity guaranteed	Obstacle avoidance	Scalable	Terrain independent
RSMN	Mobile	Yes	No	Yes	No
RELAX-RSN	Mobile	Yes	No	Yes	No
UNRES-RSMN	Mobile	Yes	No	Yes	No
DDS	Mobile	Yes	No	No	No
SEEDS	Mobile	Yes	Yes	No	No
FLOOR	Mobile	Yes	Yes	Yes	No
VOR	Mobile	No	No	Yes	No
MINIMAX	Mobile	No	No	Yes	No
PPDD	Mobile	Yes	Yes	Yes	No
CPVF	Mobile	Yes	Yes	Yes	No
VFM	Mobile	No	Yes	Yes	No
DSSA	Mobile	No	No	Yes	No
VEC	Mobile	No	No	Yes	No
CBMA	Mobile	–	–	Yes	No
GSO	Mobile	No	No	Yes	No

birds like vultures, hawks which catch their prey on the ground precisely may help in SN to reach its desired location.

4 Conclusion

Random scattering of SNs initially faces problems like environmental obstacles such as bushes, rocks, etc. and hence aerial deployment using helicopters is preferred. Even in such cases, local wind, damage due to harsh landing, etc., can affect the SN. Hence, the usage of mini parachutes or spongy landing gear is suggestable.

Various mobile sensor deployment methods have been studied. Main principles along with the advantages and disadvantages have been discussed. A comparative study has been carried out with the relocation distance as a factor.

Overall comparative analysis of various characteristics has been done and the results are discussed. As WSNs are desired to work in hazardous and unreachable locations, there is a lot of scopes for robotic engineers and researchers to develop mechanisms and algorithms to overcome the existing obstacles and to work in unideal conditions in large-scale target regions.

References

1. Nguyen, T.G., So-In, C.: Distributed deployment algorithm for barrier coverage in mobile sensor networks. Post-doctoral training program, Khon Kaen University, IEEE (2018)
2. Jun, H.E., Hongchi, S.H.I.: Constructing sensor barriers with minimum cost in wireless sensor networks. *J. Parallel Distrib. Comput.* **72**(12), 1654–1663 (2012)
3. Liu, B., Dousse, O., Wang, J., et al.: Strong barrier coverage of wireless sensor networks. In: Proceedings of the 9th ACM International Symposium on Mobile ad hoc Networking and Computing, ACM, pp. 411–420 (2008)
4. Kumar, S., Lai, T.H., Posner, M.E., et al.: Optimal sleep-wake up algorithms for barriers of wireless sensors. In: 2007 Fourth International Conference on Broadband Communications, Networks and Systems, BROADNETS, IEEE, pp. 327–336 (2007)
5. Chang, C.Y., Hsiao, C.Y., Chin, Y.T.: The k-barrier coverage mechanism in wireless mobile sensor networks. In: Proceedings of the 12th International Conference on Advances in Mobile Computing and Multimedia, ACM (2014)
6. Mostafaei, H.: Stochastic barrier coverage in wireless sensor networks based on distributed learning automata. *Comput. Commun.* **55**(1), 51–61 (2015)
7. Zou, Y., Chakrabarty, K.: Sensor deployment and target localization based on virtual forces. In: Twenty-Second Annual Joint Conference of the IEEE Computer and Communications (INFOCOM), vol. 2, no. C, pp. 1293–1303 (2003). <https://doi.org/10.1109/INFCOM.2003.1208965>
8. Tan, G., Jarvis, S.A., Kermarrec, A.M.: Connectivity-guaranteed and obstacle-adaptive deployment schemes for mobile sensor networks. *IEEE Trans. Mob. Comput.* **8**(6), 836–848 (2009). <https://doi.org/10.1109/TMC.2009.31>
9. Bartolini, N., Calamoneri, T., Fusco, E.G., Massini, A., Silvestri, S.: Autonomous deployment of self-organizing mobile sensors for a complete coverage. In: Self-Organizing Systems, pp. 194–205 (2008). https://doi.org/10.1007/978-3-540-92157-8_17

10. Kumar, A., Sharma, V., Prasad, D.: Distributed deployment scheme for homogeneous distribution of randomly deployed mobile sensor nodes in wireless sensor network. (*IJACSA*) Int. J. Adv. Comput. Sci. Appl. **4**(4), 139–146 (2013). Retrieved from www.ijacsatheai.org
11. Gupta, M., Krishna, C.R., Prasad, D.: SEEDS: scalable energy efficient deployment scheme for homogeneous wireless sensor network. In: Proceedings of the 2014 International Conference on Issues and Challenges in Intelligent Computing Techniques, ICICIT 2014, pp. 416–423 (2014). <https://doi.org/10.1109/ICICIT.2014.6781318>
12. Moreira, J.E., Rebello, V.E.F.: Proceedings of the 22nd International Symposium on Computer Architecture and High-Performance Computing, SBAC-PAD 2010: Message from the program committee co-chairs, 00(C), pp. 1597–1602 (2010). <https://doi.org/10.1109/SBAC-PAD.2010.6>
13. Lee, H.J., Kim, Y.H., Han, Y.H., Park, C.Y.: Centroid-based movement assisted sensor deployment schemes in wireless sensor networks. IEEE Veh. Technol. Conf. (2009). <https://doi.org/10.1109/VETECF.2009.5379087>
14. Liao, W.H., Kao, Y., Li, Y.S.: A sensor deployment approach using glowworm swarm optimization algorithm in wireless sensor networks. Expert Syst. Appl. **38**(10), 12180–12188 (2011). <https://doi.org/10.1016/j.eswa.2011.03.053>
15. Zhang, Y., Sun, X., Wang, B.: Efficient algorithm for k-barrier coverage based on integer linear programming. China Commun. **13**(7), 16–23 (2016). <https://doi.org/10.1109/CC.2016.7559071>
16. Ma, H., Li, D., Chen, W., Zhu, Q., Yang, H.: Energy efficient k-barrier coverage in limited mobile wireless sensor networks. Comput. Commun. **35**(14), 1749–1758 (2012). <https://doi.org/10.1016/j.comcom.2012.04.021>
17. Wang, G., Member, S., Cao, G.: Movement-assisted sensor deployment. IEEE Trans. Mob. Comput. **5**(6), 640–652 (2006)

DEA Model for Exogenously Fixed Variables Under Fuzzy Environment



Shivi Agarwal and Trilok Mathur

Abstract Data Envelopment Analysis (DEA) is developed to estimate the performance of Decision-Making Units (DMUs) and it is a nonparametric approach based on linear programming. In the real world, some inputs and outputs are not precisely known which are dealt by fuzzy DEA, where these variables could be expressed as fuzzy numbers. In addition, exogenously fixed inputs and outputs are also involved in efficiency measurement, which are not under the control of management. DEA can deal these situations with crisp variables. This paper proposed a DEA model with exogenously fixed fuzzy inputs and outputs to estimate the performance of DMUs having exogenously fixed variables in a fuzzy environment. A numerical example is also presented.

Keywords Fuzzy DEA · DMU · Exogenously fixed variables

1 Introduction

DEA is a technique based on mathematical programming to calculate the relative performance of a homogeneous DMU that includes various input and output variables. The weighted sum of the outputs (virtual output) divided by the weighted sum of inputs (virtual input) is considered as relative efficiency of a DMU. DEA models calculate the weights of input and output variables that obtain the maximum relative efficiency of a DMU and keeping the efficiency of other DMUs ≤ 1 for the same weights [1]. The assessment of efficiency by DEA can be done in two ways: by maximizing the outputs with given set of inputs named as input-oriented DEA model

S. Agarwal (✉) · T. Mathur

Department of Mathematics, BITS Pilani, Pilani 333031, Rajasthan, India
e-mail: shividma@gmail.com; shivi@pilani.bits-pilani.ac.in

T. Mathur

e-mail: mathur.trilok76@gmail.com; tmathur@pilani.bits-pilani.ac.in

or by minimizing the inputs with the set outputs named as output-oriented DEA model. Present models of DEA have limited applications to common crisp input and output variables. There are examples when qualitative data is also involved in efficiency measurement, which cannot be quantified exactly. In such cases, the data can be represented as fuzzy numbers. In addition, exogenously fixed inputs and outputs are also considered in efficiency measurement which are not under the control of management.

In this paper, DEA model with exogenously fixed variables [2] is extended in the fuzzy environment to deal with the abovementioned situation for calculating efficiency of DMUs with fuzzy exogenously fixed variables in both input-oriented and output-oriented case. There are many methods to solve fuzzy basic DEA models [3–7]. The suggested model in this paper is based on the α -cuts of the fuzzy variables. This newly developed model based on α -cut is obtained by equating two intervals, interval of the left-hand side and interval of the right-hand side of all constraints (equality/inequality).

Organization of the remaining paper is given as follows: Section 2 explains the suggested DEA model with fuzzy exogenously fixed data. A numerical example of the suggested fuzzy DEA model is presented in Sect. 3, conclusions and scope for future work are given in Sect. 4.

2 Proposed Fuzzy DEA Model with Exogenously Fixed Variables

The efficiency of the set of m DMU ($DMU_j; j = 1 \dots m$) is evaluated for the elaborate new DEA model. The efficiency of DMU_j is considered using I fuzzy inputs defined by fuzzy numbers $\tilde{X}_{aj}; a = 1, 2 \dots, I$ to obtain O fuzzy outputs defined by fuzzy numbers $\tilde{Y}_{bj}; b = 1, 2, \dots, O$. The fuzzy DEA model for input orientation with exogenously fixed inputs is given as follows:

$$\begin{aligned}
 \tilde{E}_k &= \text{Min } \theta_k \\
 &\text{subject to} \\
 &\sum_{j=1}^m \lambda_{jk} \tilde{Y}_{bj} \geq \tilde{Y}_{bk} \quad \forall b = 1, \dots, O \\
 &\sum_{j=1}^m \lambda_{jk} \tilde{X}_{aj} \leq \theta_k \tilde{X}_{ak} \quad \forall a \in I_D \\
 &\sum_{j=1}^m \lambda_{jk} \tilde{X}_{aj} \leq \tilde{X}_{ak} \quad \forall a \in I_F \\
 &\lambda_{jk} \geq 0 \quad \forall j = 1, \dots, m,
 \end{aligned} \tag{1}$$

where $\{\tilde{X}_{aj}; a \in I_D\}$ is the discretionary inputs set and $\{\tilde{X}_{aj}; a \in I_F\}$ is the exogenously fixed inputs set of j th DMU. Similarly, the output-oriented fuzzy DEA with exogenously fixed outputs is given as follows:

$$\begin{aligned}
 \tilde{E}_k &= \text{Max } \phi_k \\
 \text{subject to} \\
 \sum_{j=1}^m \lambda_{jk} \tilde{Y}_b &\geq \phi_k \tilde{Y}_{bk} \quad \forall b \in O_D \\
 \sum_{j=1}^m \lambda_{jk} \tilde{Y}_{bj} &\geq \tilde{Y}_{bk} \quad \forall b \in O_F \\
 \sum_{j=1}^m \lambda_{jk} \tilde{X}_{aj} &\leq \tilde{X}_{ak} \quad \forall a = 1, 2, \dots, I \\
 \lambda_{jk} &\geq 0 \quad \forall a = 1, \dots, m,
 \end{aligned} \tag{2}$$

where $\{\tilde{Y}_{bj}; b \in O_D\}$ is discretionary outputs set and $\{\tilde{Y}_{bj}; b \in O_F\}$ is the exogenously fixed outputs set of j th DMU.

To solve the above fuzzy DEA model, we extend the λ -cut method based on Zadeh's extension principle [8], developed by Agarwal [9] for both input-oriented and output-oriented DEA model for fuzzy exogenously fixed input and output variables, respectively. In this method, the inputs and outputs could be expressed as a closed crisp interval $[(X_{aj})_\lambda^L, (X_{aj})_\lambda^U]$ and $[(Y_{bj})_\lambda^L, (Y_{bj})_\lambda^U]$, respectively, at different λ -cuts. Thus, the fuzzy model of DEA is remodeled for the collection of DEA parametric models with different λ -cuts where $\lambda \in [0, 1]$ is a parameter. The resulting model can be solved as an LP model for a preset parameter λ to observe how the efficiency scores of the DMUs change according to the change of the level of the confidence interval. The lower and the upper bound of λ -cut of the fuzzy efficiency of DEA model will become two-level linear model that could be simplified by the traditional linear model based on the concept suggested by Liu and Chuang [6]. The lower and the upper bounds of fuzzy efficiency obtained by input-oriented DEA model with exogenously fixed inputs are given by

$$\begin{aligned}
 (E_k^I)_\lambda^L &= \text{Min } \theta_k \\
 \text{subject to} \\
 \lambda_{jk}(Y_{bk})_\lambda^L + \sum_{j=1, j \neq k}^m \lambda_{jk}(Y_{bj})_\lambda^U &\geq (Y_{bk})_\lambda^L \quad \forall b = 1, \dots, O \\
 \lambda_{jk}(X_{ak})_\lambda^U + \sum_{j=1, j \neq k}^m \lambda_{jk}(X_{aj})_\lambda^L &\leq \theta_k(X_{ak})_\lambda^U \quad \forall a \in I_D \\
 \lambda_{jk}(X_{ak})_\lambda^U + \sum_{j=1, j \neq k}^m \lambda_{jk}(X_{aj})_\lambda^L &\leq (X_{ak})_\lambda^U \quad \forall a \in I_F \\
 \lambda_{jk} &\geq 0 \quad \forall j = 1, \dots, m,
 \end{aligned} \tag{3a}$$

and

$$\begin{aligned}
 & (E_k^I)_\lambda^U = \text{Min } \theta_k \\
 & \text{subject to} \\
 & \lambda_{jk}(Y_{bj})_\lambda^U + \sum_{j=1, j \neq k}^m \lambda_{jk}(Y_{bj})_\lambda^L \geq (Y_{bk})_\lambda^U \quad \forall b = 1, \dots, O \\
 & \lambda_{jk}(X_{aj})_\lambda^L + \sum_{j=1, j \neq k}^m \lambda_{jk}(X_{aj})_\lambda^U \leq \theta_k(X_{ak})_\lambda^L \quad \forall a \in I_D \\
 & \lambda_{jk}(X_{aj})_\lambda^L + \sum_{j=1, j \neq k}^m \lambda_{jk}(X_{aj})_\lambda^U \leq (X_{ak})_\lambda^L \quad \forall a \in I_F \\
 & \lambda_{jk} \geq 0 \quad \forall j = 1, \dots, m.
 \end{aligned} \tag{3b}$$

These crisp linear programming problems will become parametric programming with parameter λ . The fuzzy efficiency can be obtained by inverting $(E_k)_\lambda^L$ and $(E_k)_\lambda^U$ with respect to λ . Furthermore, the set of intervals $\{[(E_k)_\lambda^L, (E_k)_\lambda^U] | \lambda \in (0, 1]\}$ reveals the curve of fuzzy efficiency. Similarly, we can obtain the pair of the crisp parametric programming model of output-oriented fuzzy DEA model with exogenously fixed outputs.

$$\begin{aligned}
 & (E_k^O)_\lambda^L = \text{Max } \phi_k \\
 & \text{subject to} \\
 & \lambda_{jk}(Y_{bk})_\lambda^L + \sum_{j=1, j \neq k}^m \lambda_{jk}(Y_{bj})_\lambda^U \geq \phi_k(Y_{bk})_\lambda^L \quad \forall b \in O_D \\
 & \lambda_{jk}(Y_{bk})_\lambda^L + \sum_{j=1, j \neq k}^m \lambda_{jk}(Y_{bj})_\lambda^U \geq (Y_{bk})_\lambda^L \quad \forall b \in O_F \\
 & \lambda_{jk}(X_{ak})_\lambda^U + \sum_{j=1, j \neq k}^m \lambda_{jk}(X_{aj})_\lambda^L \leq (X_{ak})_\lambda^U \quad \forall a = 1, 2, \dots, I \\
 & \lambda_{jk} \geq 0 \quad \forall j = 1, \dots, m,
 \end{aligned} \tag{4a}$$

and

$$\begin{aligned}
 & (E_k^O)_\lambda^U = \text{Max } \phi_k \\
 & \text{subject to} \\
 & \lambda_{jk}(Y_{bj})_\lambda^U + \sum_{j=1, j \neq k}^m \lambda_{jk}(Y_{bj})_\lambda^L \geq \phi_k(Y_{bk})_\lambda^U \quad \forall b \in O_D \\
 & \lambda_{jk}(Y_{bj})_\lambda^U + \sum_{j=1, j \neq k}^m \lambda_{jk}(Y_{bj})_\lambda^L \geq (Y_{bk})_\lambda^U \quad \forall b \in O_F \\
 & \lambda_{jk}(X_{aj})_\lambda^L + \sum_{j=1, j \neq k}^m \lambda_{jk}(X_{aj})_\lambda^U \leq (X_{ak})_\lambda^L \quad \forall a = 1, \dots, I \\
 & \lambda_{jk} \geq 0 \quad \forall j = 1, \dots, m.
 \end{aligned} \tag{4b}$$

The ranking of DMUs by fuzzy DEA model is not direct as the efficiency score of a DMU obtained by fuzzy DEA model is a fuzzy number. In this study, we have used λ -cut-based method for ranking the DMUs given by Chen and Klein [10]. The ranking index for the j th DMU as

$$I_j = \frac{\sum_{k=0}^s ((E_j)_{\lambda_k}^U - \eta)}{\left[\sum_{k=0}^s ((E_j)_{\lambda_k}^U - \eta) - \sum_{k=0}^s ((E_j)_{\lambda_k}^L - \xi) \right]}; s \rightarrow \infty \quad (5)$$

$$\eta = \min_{k,j} \left\{ (E_{jk})_{\lambda_k}^L \right\}; \quad \xi = \max_{k,j} \left\{ (E_{jk})_{\lambda_k}^U \right\},$$

where s is λ -cuts and $\lambda_k = kt/s$, $k = 0, \dots, s$, t is the height of efficiency scores. The higher score of index I_j of j th DMU provides the information that DMU_j is more efficient. The index I_j is a crisp number that lies in $[0, 1]$, which can be compared with the efficiency scores obtained by crisp DEA model. The requirement for the validation of the method is s tends to infinite for but the method works practically only for s is equal to 3 or 4.

3 Numerical Demonstration

To demonstrate the suggested fuzzy output-oriented DEA model with exogenously fixed output, consider a hypothetical data of single input and two outputs out of which one is discretionary output and one is exogenously fixed output (can't be controlled by management) for four DMUs. Data is shown in Table 1.

Table 2 depicts the results of the interval of the efficiency scores of distinct eleven values of λ ($0, 0.1, \dots, 1.00$) with the step of 0.1 and ranking index of four DMUs.

The results evince that all DMUs have the fuzzy efficiency scores, which is in the form of a trapezoidal fuzzy number. The efficiency score at $\lambda = 0$ depicts the maximum range where at $\lambda = 1$ produces a crisp real number.

Next, we calculate efficiency indices as $I_{D_1} = 0.90$, $I_{D_2} = 0.71$, $I_{D_3} = 0.54$, $I_{D_4} = 0.51$, which shows that $I_{D_1} > I_{D_2} > I_{D_3} > I_{D_4}$, which further shows that DMU D₁ is the most efficient and DMU D₄ is least efficient.

Table 1 Data of DMUs

DMUs	D ₁	D ₂	D ₃	D ₄
Input 1	2	3	4	2.5
Output 1 (D)	(0.5, 0.75, 1)	(0.75, 1, 1)	(0.5, 0.75, 1)	(0.25, 0.5, 0.75)
Output 2 (F)	(0.25, 0.5, 0.75)	(0.5, 0.75, 1)	(0.75, 1, 1)	(0.25, 0.5, 0.75)

Table 2 Efficiency scores

λ values	Efficiency scores			
	D ₁	D ₂	D ₃	D ₄
0.0	[0.75, 1]	[0.50, 1]	[0.25, 1]	[0.2, 1]
0.1	[0.79, 1]	[0.53, 1]	[0.27, 1]	[0.23, 1]
0.2	[0.83, 1]	[0.56, 1]	[0.29, 1]	[0.25, 1]
0.3	[0.86, 1]	[0.59, 1]	[0.31, 1]	[0.28, 1]
0.4	[0.90, 1]	[0.63, 1]	[0.33, 1]	[0.31, 1]
0.5	[0.94, 1]	[0.67, 1]	[0.36, 1]	[0.34, 1]
0.6	[0.98, 1]	[0.71, 1]	[0.38, 1]	[0.38, 1]
0.7	[1, 1]	[0.75, 1]	[0.41, 1]	[0.41, 1]
0.8	[1, 1]	[0.79, 1]	[0.44, 1]	[0.45, 0.63]
0.9	[1, 1]	[0.84, 1]	[0.47, 1]	[0.49, 0.58]
1.0	[1, 1]	[0.89, 0.89]	[0.5, 0.5]	[0.53, 0.53]
Index I	0.90	0.71	0.54	0.51

4 Conclusion

DEA models applying to measure the relative efficiency of DMUs with various input and output variables. This paper suggested a fuzzy model of DEA with exogenously fixed input and output variables to assess the relative efficiency of DMUs in a fuzzy environment based on Zadeh's extension principle with λ -cut technique. The proposed DEA model measuring the efficiency of DMUs involves the variables which are not controlled by management in a fuzzy environment in more objective form. The ranking of DMUs is done by ranking method given by Chen and Klein. Further, a numerical example is given to demonstrate the proposed fuzzy DEA model with exogenously fixed variables. The efficiency obtained are fuzzy numbers, which provides more information for management. The extension of fuzzy DEA technique has become more useful for the application.

References

- Charnes, A., Cooper, W.W., Rhodes, E.: Measuring the efficiency of decision making units. *Eur. J. Oper. Res.* **2**, 429–441 (1978)
- Banker, R.D., Morey, R.: Efficiency analysis for exogenously fixed inputs and outputs. *Oper. Res.* **34**, 513–521 (1986)
- Kao, C., Liu, S.T.: Fuzzy efficiency measure in data envelopment analysis. *Fuzzy Sets Syst.* **113**, 427–437 (2000)
- Wen, M., Li, H.: Fuzzy data envelopment analysis (DEA): model and ranking method. *J. Comput. Appl. Math.* **223**(2), 872–878 (2009)
- Guo, P., Tanaka, H.: Fuzzy DEA: a perceptual evaluation method. *Fuzzy Sets Syst.* **119**, 149–160 (2001)

6. Liu, S.T., Chuang, M.: Fuzzy efficiency measure in fuzzy DEA/AR with application to university libraries. *Fuzzy Sets Syst.* **36**(2), 1105–1113 (2009)
7. Agarwal, S.: SBM data envelopment analysis in fuzzy environment. *Math. Sci. Int. Res. J.* **3**(1), 478–484 (2014)
8. Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets Syst.* **1**, 3–28 (1978)
9. Agarwal, S.: Efficiency measure by fuzzy data envelopment analysis model. *Fuzzy Inf. Eng.* **6**, 59–70 (2014)
10. Chen, C.B., Klein, C.M.: A simple approach to ranking a group of aggregated fuzzy utilities. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **27**, 26–35 (1997)

Status of Small Hydropower in Himachal Pradesh: A Review



**Kamal Kashyap, Robin Thakur, Anil Kumar, Raj Kumar Saini
and Amar Raj Singh Suri**

Abstract From many years small-scale hydro control plants have been inspected by different individuals all through the world because of its benefits of offering preferred execution over the ordinary petroleum derivatives to meet the demand of energy. Thinking about the arrangement of energy, the change of hydro power will share in diminishing of ozone depleting substance discharge and better flexibility in the grid activity. India is searching forward for discovering new potential outcomes in sustainable power source. With lot of glaciers and Rivers, western Himalayan States like J&K, Uttarakhand and Himachal Pradesh has a lot of potential in Hydro Power. Himachal Pradesh with the aggregate geological area of 55673 sq km is the eighteenth state of Union of India. Himachal Pradesh with an exceptionally complicated Geography, Regions of the state shifts from high mountains in east to planes in north west. With five rivers named Satluj, Beas, Yamuna, Ravi and Chenab Himachal Pradesh have the aggregate capability of more than 27000 MW in hydro power. Since most power plants now-a-day's utilize large turbines for the low power generation leads to losses and enhances the overall cost, thus this study will be useful to decrease the expense of the plant. This paper mainly focus on past, present and future situation of small-scale hydro extends in Himachal Pradesh. Beside this the paper demonstrates the potential outcomes of hydropower in State.

K. Kashyap

Department of Mechanical Engineering, A P Goyal Shimla University, Shimla, HP, India
e-mail: kamal.maximus23@gmail.com

R. Thakur (✉) · A. Kumar · R. K. Saini · A. R. S. Suri

School of Mechanical and Civil Engineering, Shoolini University, Solan 17329, Himachal Pradesh, India
e-mail: robinsampark.thakur@gmail.com

A. Kumar

e-mail: anilkumar88242@gmail.com

R. K. Saini

e-mail: rajsaini.acet@gmail.com

A. R. S. Suri

e-mail: assuri@gmail.com

Keywords Energy · Small hydro power

1 Introduction

The power deficiency is a noteworthy obstacle India needed to confront. For the nation's economy to support, power is its ancient rarity. A solitary electric power production [1] utilizing inexhaustible assets is one among the most required and beneficial strategy to create power. Particularly in places slacking grid connection, for example, rural side of the nation small scale hydro control is a vital asset [2–4] in which power is produced by changing over the potential energy of water which remains as the purest type of energy on the planet. The water accordingly produced is likewise kept for water system [5–8] and other residential purposes after age of power. The idea of utilizing running waters to produce power was started by the presentation of water wheel in Wisconsin in 1882 at the Fox stream. In this century hydro control had denoted a place in electrical energy creation around the globe. The power age different from plant to plant contingent upon a few angles and those plants which create power lesser than 100 kW are named as miniaturized scale hydro control plants. These small hydro plants devour less space, solid and cost-effective then the petroleum derivatives [9]. Because of its notable highlights it pays a way in foundation and advancement of micro hydro plants in the nation as rather than mega hydro projects that occupies more area [10].

There are tremendous number of sites with high potential and more noteworthy interest for electricity have been identified as low head pico assets and utilization of such potential destinations is essential [11, 12]. A few of the pico hydro plants are situated in remote territories or in forest areas [13].

2 Past of Hydropower in Himachal Pradesh

Prior to the autonomy the establishment of the Hydro Power Projects was restricted to just with some princely states having better budgetary assets. The historical backdrop of Hydro power extends in Himachal Pradesh demonstrates that the working in this field had started in 1908 in Chamba at that point, the Hilly state. Shri Bhuri Singh, the then ruler (prince) of Chamba state installed first hydro power project in the state and in this manner cleared an approach to other states to introduce comparable hydro power projects in their states. The leaders of the other regal states followed him and began to generate power out of sourced water through projects. On March 10, 1933 WILLINGDONE then viceroy and governor general of British India inaugurated and handed over the BASSISHANAN Hydro Power project to the people of Mandi of Kangra State at that time. In 1948 the aggregate production of electricity was 550 kW in Himachal Pradesh which demonstrates that this facility was constrained to just a few people in Himachal Pradesh. The government of Himachal Pradesh

understood the need and utility of Hydro Projects to improve the financial and all-round development of the state.

Hence, the HP Government passed the Power Legislation Bill in HP Legislative Assembly in 1948 and took over the control of hydro power projects from private sector. In 1971 the State Electricity Board was setup to control and watch all the activities related to generation, transmission and distribution of hydro power produced in different projects. This progression of the legislature expanded the workload of HPSEB having antagonistic impact in the productivity and working of Hydro Power Projects. An endeavor made to limit the workload of HPSEB and the three separate wings were set up to make the HPSEB more efficient in 2010.

The wings and their work are as follows.

2.1 Generation

The purpose of this wing was to control every one of the exercises identified with the generation of Hydro Power. The work in such manner was appointed to HPPCL (Himachal Pradesh Power Corporation Company Ltd.).

2.2 Transmission

The work and control of Transmission of power was appointed to HPTCL (Himachal Pradesh Power Transmission Company Ltd.).

2.3 Distribution

At long last the work in regard to circulation was allocated to HPSEB (Himachal Pradesh State Electricity Board Ltd.) so as to control and organize the exercises of various wings. The Directorate of Power was setup in 2010.

3 Present

3.1 Hydropower in Himachal Pradesh Current Scenario

There is a significant potential of hydropower in Himachal Pradesh. Numerous agencies are doing work in hydro with the aggregate potential of 27436 MW. There are around 1000 projects under various Hydropower agencies in Himachal Pradesh (both state possessed and private). Currently, the total commissioned projects are 148 that

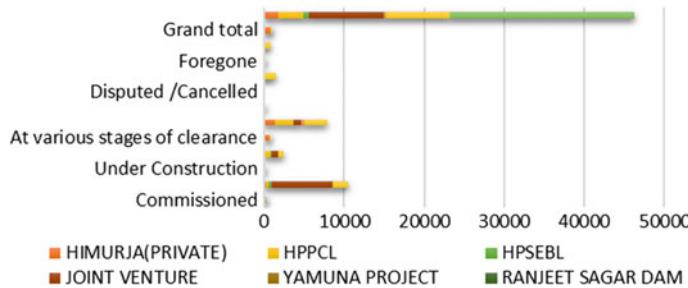


Fig. 1 Hydropower agencies in Himachal Pradesh generation in (MW) [14]

Table 1 Categorization of different hydro power plants in India [22]

Category	Description
Mini	Less than 5 kW of electricity
Micro	Up to 100 kW
Small	Up to 25 MW
Large	Above 25 MW up to 10 GW

records for the generation of 10534 MW of power by 8 different agencies as shown in Fig. 1.

State and Centre with the joint venture are creating greatest of 71% with the aggregate efficiency of 7457.73 MW by 12 ventures. Assisted potential outcomes in joint venture are greatest as 1 project of 800 MW is under development and 4 projects of 456 MW are at various phases of clearances [14, 15]. Private division is likewise booming up as it produces the aggregate of 1964.9 MW of power. Private with the aggregate of 22 commissioned, 21 under development and 52 at various phases of freedom holds the second biggest share. Other agencies like HPSEBL are creating 487.55 MW of electricity, Himurja (Private) 297.75 MW, Himurja (STATE) 2.37 MW. HPPCL is creating 165 MW and Ranjeet Sagar Dam and Yamuna projects are producing 27.6 and 131.57 MW of power [14].

Hydroelectric power plants are sorted into four distinct classifications. There is no particular definition to characterize the diverse sizes of hydropower [16]. A large portion of the nations consider 10 MW establishment limit as Small-Scale Plant. In Asian nations like China the establishment limit up to 50 MW is considered as SHP. In India, this limit is 25 MW. In European nations like Canada and USA, the farthest point is extended to again 25 and 30 MW. India categorization framework to various power plants are showed in Table 1.

3.2 *Commissioned Basin Wise Distribution of Hydropower*

The five rivers Satluj, Beas, Chenab, Yamuna and Ravi are the life line of Himachal Pradesh. Specifically, or in a roundabout way every individual in the state is dependent to these rivers. All the enormous towns Rampur, Chamba, Bilaspur, Kullu and Nahan are situated on the bank of these rivers. These rivers are in-charge of income for the state as the state is reliant on vegetation, Hydropower and tourism for income [17]. The state delivers the aggregate of 0.9 million tons of vegetable every year over 70–80% of the total area of 10 districts are under fruits and vegetables and rest 2 districts over 20–30% of total area fruits and vegetables are grown [18] (Fig. 2).

Hydropower with the aggregate age of 10254 MW just 3% is delivered by Small, Mini and Micro hydro projects staying 97% is produced from medium and large power plants. Satluj basin with the aggregate of 19 appointed projects creates the aggregate of 4494.25 MW of power which conveys the most extreme level of 57% of the aggregate power generation of state as shown if Fig. 3. All the major ventures like NJPC 1500 MW, Bhakra 1478.73 MW, KolDam 800 MW and Rampur 412 MW are situated in the bank of stream Satluj as shown in Table 2.

Beas creates the aggregate of 2681.9 MW of power and is the second biggest supplier of power in the state. Beas–Satluj link 990 MW, Parbati-3 520 MW and Pong Dam in region of Mandi 396 MW are situated in the basin of river Beas. Nine small scale projects are additionally appointed on Beas basin with the aggregate of 95.9 MW of power. Ravi River creates the aggregate of 1374 MW of power out of which 1357 MW are produced by large scale and only 17.45 MW are created by small scale projects. Yamuna produces the aggregate of 79.95 MW of power by both large and small-scale projects. Chenab with the greatest flow of water is creating

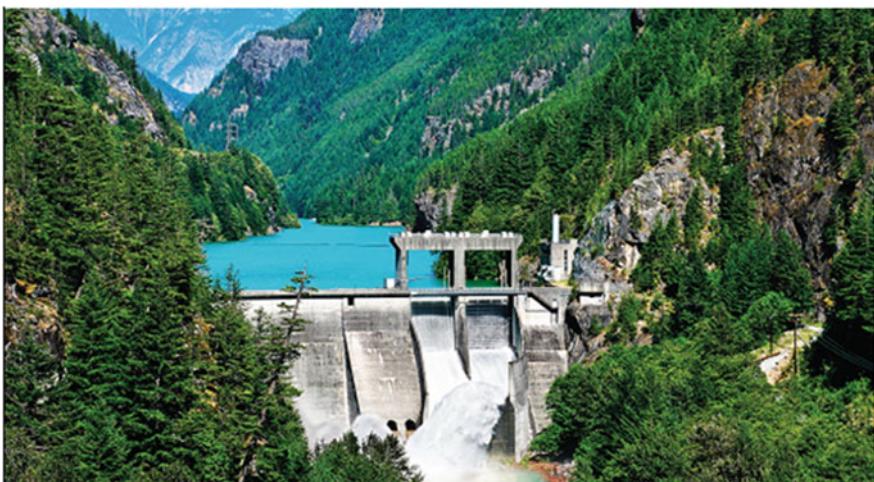


Fig. 2 Karcham Dam district Kinaur Himachal Pradesh

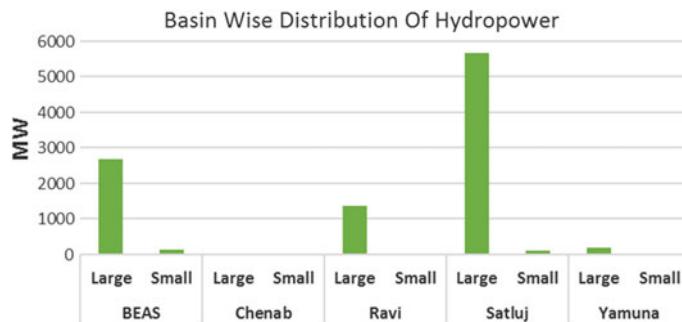


Fig. 3 Small/Large-scale distribution river wise in (MW) [21]

Table 2 Some large-scale power projects in Himachal Pradesh [18]

Sr. No	Name of Dam	State/ Centre	Installation capacity	Height of Dam (m)	District	Basin
1	NJPC	Centre/ Joint	1500	62.5	Kinnaur/ Shimla	Satluj
2	Bhakra	Centre	1325	225.55	Bilaspur	Satluj
3	Chamera- I	Centre	540	140	Chamba	Ravi
4	KolDam	Centre	800	167	Mandi/ Bilaspur	Satluj
5	Pong Dam	Centre	360	133	Kangra	Beas
6	Pandoh	Centre	990	76	Mandi	Beas
7	Parbati-3	Centre	520	43	Kullu	Beas
8	Bairasiul	Centre	180	53	Chamba	Ravi

the minimum of power as there is just 3 small scale projects commissioned with the aggregate of 4.9 MW of power.

4 Status of Small Hydropower Basin Wise

Numerous civilizations were born on the bank of rivers. The river banks have always stayed holy place for the people of India. Rivers like Ganges, Yamuna, Kaveri, Saraswathi, Brahmaputra and Sindhu are associated with their incredible awe and reverence for that. The river arrangement of Himachal is exceptionally confused. The principle rivers of Himachal Pradesh like Satluj, Yamuna, Beas chenab and Ravi are made by Glaciers and Lakes. In rainy season they are flooded and conveys 70% of the river water into the ocean. Hydropower ventures have been set up in the basin of these rivers. These undertakings are playing an extraordinary and imperative part in making the state prosperous. Around 1500 glaciers exist in Himalayas which give

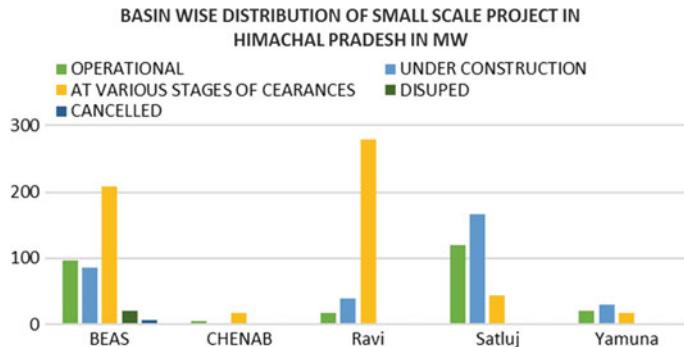


Fig. 4 Basin-wise distribution of small scale project in Himachal Pradesh in MW [19]

a great amount of water to these rivers. Barashigri is the biggest glacier it is 25 km long and 3 km wide is the centralized server of Chenab river (Fig. 4).

Different glaciers like Chandra glaciers, Bhaga glaciers, the Lady of Kelong glaciers etc. are giving an awesome part in the advancement of these rivers. 18% of the Himalayas constantly secured with snow and assumes an essential part to provide water for these rivers. These waterways convey an awesome potential for Small Scale hydropower Projects as the total of 500 low head turbines can be installed on different places of state [19].

5 Difficulties for Small Hydro Projects in Himachal Pradesh

As narrated above Himachal Pradesh is rich in the field of natural resources such as water which is the main source of life. It has been observed that not only in Himachal Pradesh but in the whole world the water level has gone down. Therefore, it is said that the third world war will be fought for water due to scarcity of water which has been caused due to uncertainty of rains and disappearing of the glaciers because of global warming. However, Himachal Pradesh is rich in the field of water. Besides rains, different rivers and fountains are the main source of water in it. Many Hydro power projects have been installed in the government and private sector on the basis of important rivers of Himachal Pradesh such as Satluj, Beas, Yamuna, Ravi and Chenab. Huge amount is being spent on these Hydro power projects in both the sectors. But it is a matter of great concern that the projects so installed have failed to achieve the targets that were expected to be achieved. It has impeded the rapid progress of the state in the field of projects. Some factors and challenges are responsible for it. If we investigate the matter prudently we will find that two sorts of elements have become obstacles in the progress and smooth functioning of the Hydro power projects.

5.1 Natural Factors

Most of the projects are installed in distant and remote places of state. It is very difficult to carry the huge machinery and other necessary equipment to the site of the projects due to unavailability of sufficient roads and its tough geography. Beside this, problem of experts and skilled labor is faced by most of the projects as this is very difficult for contractors to find skilled labor in these distant places. The farmers also reside on the banks of rivers and perform their agricultural work in their fields. While constructing a project sufficient land is required to store water in big dams. The farmers are not interested to sale their land for project purposes as it is very difficult for them to earn their living without agricultural land. The hills and mountains of the state are becoming barren due to unwise cutting of plants and trees resulting in the increase of global warming. A large number of plants and trees are removed while the work for the installation of the project is going on. In this way some kinds of shrubs and plants will be destroyed totally and eliminated from the word scene the vibration caused due to heavy machinery sometimes make cracks in the hills which results in the landslide and the elimination of our spontaneous mountains.

5.2 Human Factors

Most of big cities in Himachal Pradesh are established in the bank of rivers in case of failure of dam there may cause a huge amount of human life loss, so this become very difficult for companies to get NOC from local bodies for land accusation for the construction of the dam. A huge amount of population gets affected at the time of land accusation doing rehabilitation of thousands of families is very difficult for government. Methane emotion is also a major factor as it is noted that dam reservoirs emit a large amount of methane this a cause a serious breathing problem to the people nearby [20].

5.3 Spiritual Factors

Himachal Pradesh is known as the land of gods. Most of the dam sites are in very remote and tribal areas. People in most of the rural and tribal areas are very much devoted to their local gods and goddesses. To construct dams in these remote and tribal areas may cause tempering with people's faith so this is very difficult for hydropower developers to get NOC from Grama Panchayats for construction and land accusation.

6 Scope of Small Hydro Projects

Water mills (Gharats) are the used by the Himalayan people from centuries to grind grains and for irrigation. "Gharats" act as the source of life hood for the local population of hilly terrain. "Gharats" are the places where the water from different perineal springs are utilized by local population to operate the traditional water mills. Gharats has no effect on the ecosystem and have a low maintenance cost due to the locally used materials and technologies. Now electrically operated grinding mills are used in place of Gharats due to this many Gharats has shut down due to the low income. The locations of these closed Gharats could be the answer for the energy demand of the state. These locations are very useful as we find that the head of water is enough to install a small hydroelectric unit and on the another end the hydro unit thus installed will not affect the eco system. State government with help from Centre government has taken measures to use this unseen treasure for electricity production in remote areas.

7 Conclusion

On the basis of different aspects discussed in this research paper it is concluded that Himachal Pradesh has covered a long way from the day Shri Bhuri Singh first started hydro power project in Chamba. Lot of power is harnessed by the different companies doing work in the field of hydro power but a lot of potential is still untrapped. Over 97% of the total hydro power is generated from large and medium projects but small hydro projects play an important role to fulfill the energy needs of people in remote areas. Small hydro projects right now with the total of just 3% are increasing rapidly as many projects are under construction and clearances phases. As we know small hydro projects are very important for rural development but companies making hydropower face many problems on human, administration and spiritual level. Water mills locally said "gharats" could be the future of small hydro project in state because with changes they can generate electricity without interrupting the ecosystem.

References

1. Nasir, B.A.: Design of micro hydro electric power station. *Int. J Eng. Adv. Technol.* **2**(5), 39–47 (2013)
2. Varughese, A., Michael, P.A.: Electrical characteristics of micro-hydro power plant proposed in Valara waterfall. *Int. J. Innov. Technol. Explor. Eng.* **2**(2), 128–131 (2013)
3. Archetti, R.: Micro hydro electric power: feasibility of a domestic plant. *Procedia Eng.* **21**, 8–15 (2011)
4. Yannopoulos, S.I., Lyberatos, G., Theodossiou, N., Li, W., Valipour, M., Tamburrino, A., et al.: Evolution of water lifting devices (pumps) over the centuries worldwide. *Water* **7**(9), 5031–5060 (2015)

5. Valipour, M.: A comprehensive study on irrigation management in Asia and Oceania. *Arch. Agron. Soil Sci.* **61**(9), 1247–1271 (2015)
6. Valipour, M.: Future of agricultural water management in Africa. *Arch. Agron. Soil Sci.* **61**(7), 907–927 (2015)
7. Valipour, M.: Sprinkle and trickle irrigation system design using tapered pipes for pressure loss adjusting. *J. Agric. Sci.* **4**(12), 125–133 (2012)
8. Valipour, M., Sefidkouhi, M.A.G., Eslamian, S.: Surface irrigation simulation models: a review. *Int. J. Hydrol. Sci. Technol.* **5**(1), 51–70 (2015)
9. Mohibullah, M., Radzi, A.M., Hakim, M.I.A.: Basic design aspects of micro hydro power plant and its potential development in Malaysia. In: Proceedings of Power Energy 2004
10. Madanhire, I., Gudukeya, L.: Efficiency improvement of pelton wheel and cross flow turbines in micro-hydro power plants: case study. *Int. J. Eng. Comput. Sci.* **2**(2), 416–432 (2013)
11. Williamson, S.J., Stark, B.H., Booker, J.D.: Low head pico hydro turbine selection using a multi-criteria analysis. *Renew. Energy* **61**, 43–50 (2014)
12. Haidar, A.M.A., Senan, M.F.M., Noman, A., Radman, T.: Utilization of Pico hydro generation in domestic and commercial loads. *Renew. Sustain. Energy Rev.* **16**(1), 518–524 (2012)
13. Ho-Yan, B.P.: Design of a low head pico hydro turbine for rural electrification in Cameroon. Canada: a thesis presented to The University of Guelph (2012)
14. Energy, D.: Detail Projects Report. Directorate of Energy, Shimla (2018)
15. Energy, D.: Status of Hydropotential in Himachal Pradesh. Directorate of Energy, Shimla (2018)
16. Adu, D.: A case study of status and potential of small hydro-power plants in Southern African development community. *Energy Procedia* 352–359 (2017)
17. Bhardwaj, M.: Directory of water resources in Himachal Pradesh. HP State Council for Science Technology and Environment, Shimla (2014)
18. Patikari Hydro Mohan Everest State: Basin wise detail of hydropower in Himachal Pradesh. HPSEB, Shimla (2018)
19. Climate, S.C.: Glaciers of Himachal Himalaya. <http://www.hpcce.gov.in/glaciers.aspx> (2018). Accessed 12 Apr 2018
20. Slaryia, M.: Hydroelectric power generation: Himachal Pradesh perspective. *EIJMMS* **3**(5), 193–205 (2013)
21. Himachal Pradesh Power. Economic Survey. http://admis.hp.nic.in/himachal/economics/en_power.html (2018). Accessed 3 May 2018
22. Razan, J.I.: A comprehensive study of micro-hydropower plant and its potential in Bangladesh. *ISRN Renew. Energy* 1–10 (2012)

Implementation of HMI Using VISUAL BASIC .NET for Induction Hardening Process



Himanshu Sharma, Roushan Kumar, Adesh Kumar, Mukul Kumar Gupta and Vivek Kaundal

Abstract In Industrial Automation, the use of HMI for the monitoring of process nowadays has increased with the increment in competition. But due to the high cost, some industries cannot afford it. So to make it affordable development of HMI using Visual Basics .NET has been introduced. Traditional HMI provides many functionalities, but at a higher cost. But with Visual Basic. NET we can afford the same task in manner. The designed view of VB based HMI is in proportion with a view of the automation process in real time, saves troubleshooting time of faults and helps in safety purpose. Small-scale industries now can implement HMI as a major part of automation to visualize processes without giving others life at the dangerous environment and can detect faults easily. This paper represents an implementation of HMI of induction hardening process using VB.NET.

Keywords HMI · Automation technology · Visual Basic · Pneumatic actuator · PLC

1 Introduction

Induction hardening is the process of hardening of the metal using the process of electromagnetic induction. This process is followed by quenching with water or oil to make a metal hard. Induction hardening is the noncontact-type heating process

H. Sharma (✉) · A. Kumar

Department of Electrical and Electronics, School of Engineering, University of Petroleum and Energy Studies, Dehradun, India
e-mail: thanks.himanshu@gmail.com

R. Kumar

Department of Mechanical Engineering, School of Engineering, University of Petroleum and Energy Studies Dehradun, Dehradun, India

M. K. Gupta · V. Kaundal

Department of Electrical and Electronics Engineering, University of Petroleum and Energy Studies, Dehradun, India

of metal. In this, the job is placed between the coils, and when AC current flows from the coil a magnetic field is developed which penetrates the job and leads to the development of eddy current inside it, which causes metal heating. Industrial automation is the process of replacing human labor with the use of controls over the machine. There is an overall process of automation.

- (1) We need a brain to think and to make a sequence of processes like we do while performing tasks.
- (2) Other than this we need a feedback from actuators that work is completed we can proceed to next step.

The introduction of automation in industries improved productivity with reduction in time. PLC stands for Programmable Logic Controller, which serves as a brain of a machine here, which takes care of sequences of a process as well as control the process. PLC was introduced in industries to replace hard-wired devices, as to program hard-wired devices for a small task it requires lots of wires which proves it time taking and tedious task.

PLC is an optocoupler device, which works on a relay logic and performs the task according to it. Rewriting of the program is easier and not tedious to do which makes it time relaxing and this is how PLC eliminates hard-wire type process in industries. To get feedback from actuators, we use sensors.

Human–Machine Interface (HMI) is an interface between human and machine by providing a visual and graphical representation of a process in real time. HMI serves as an important key to having a look on the process without physical presence of any individual. It helps in the control and supervision of process.

Visual Basic .NET is an object-oriented programming language by Microsoft using .NET Framework. Development of HMI using Visual Basics .NET is benefited in many ways . It is used worldwide, run time fee is negligible and provide security (Fig. 1).

2 Hardware Design

Conveyor section: Conveyor section comprises three sensors for the job pieces to reach the processing section. First sensor indicates the pneumatic piston to stop the job to enter in the processing section, 2nd sensor indicates another pneumatic piston to make difference between the next going job (which is now stopped with the first actuator) and remaining jobs from job array (Fig. 2) and third sensor indicates these two sensors to work according to it. This sensor is used to indicate that at this time we have a job to pick so don't eliminate the next job until this it is picked.

Heating section: After section first, when the job is picked there is 1 rotating bed on which 4 circular beds and 4 proximity switches are placed, So when the job is picked and placed on the first bed, PROXIMITY 1 is ON and rotating bed rotates to 90°. At the second position when PROXIMITY 2 is ON, the clamper clamps the job and around the bed there is coil for induction hardening and the bed is rotating, after this when PROXIMITY 3 detects job quenching process which will proceed and in

Fig. 1 Flowchart of a conveyor section

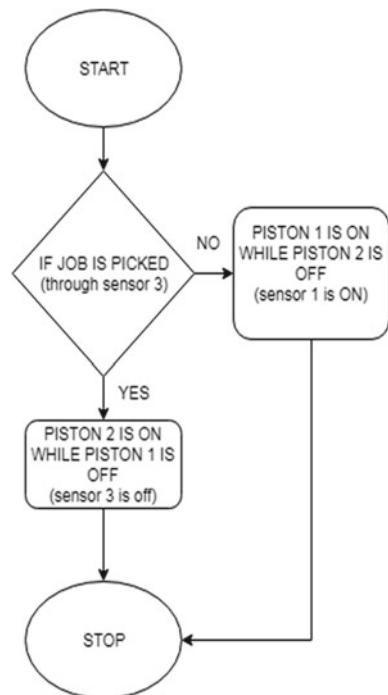


Fig. 2 Top view of a conveyor section

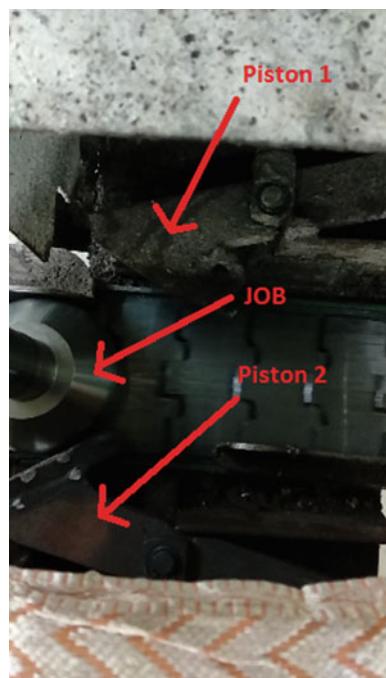
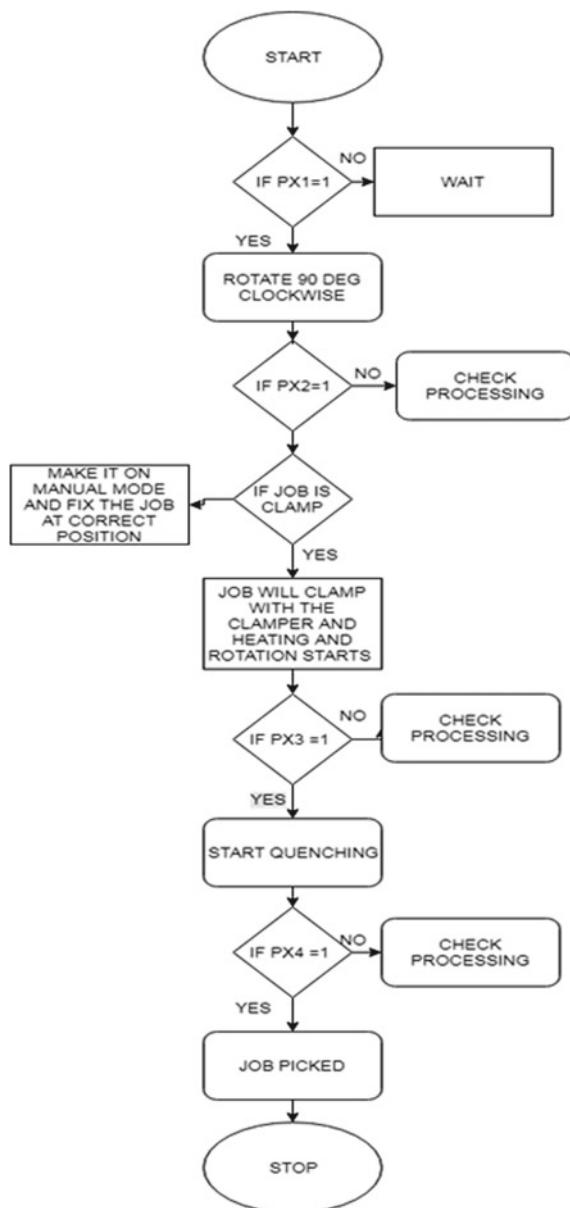


Fig. 3 Flowchart of an EFD heating



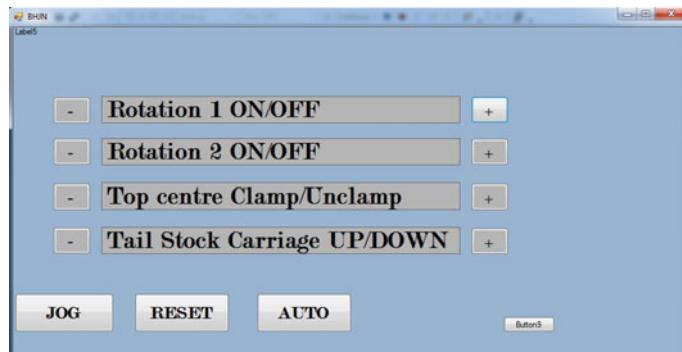


Fig. 4 HMI for heating and quenching

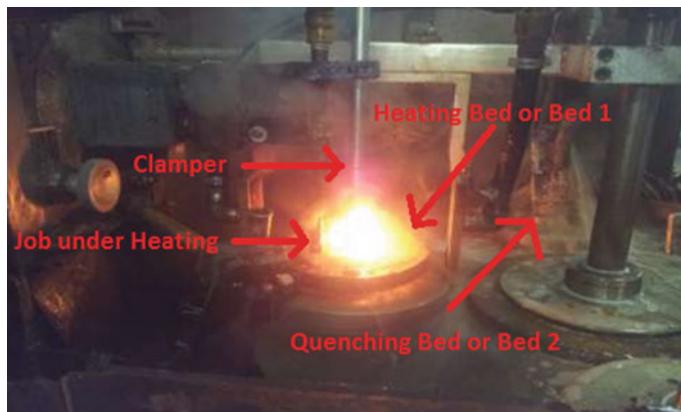


Fig. 5 Front view of a heating process

the last job when PROXIMITY 4 detects job which will be picked to next processing (Fig. 3).

HMI designing: To design an HMI, we should know the individual working of every actuator, every sensor, etc. HMI provides an interface between human and machine so for every machine there is an HMI to provide real-time processing and the user can give input in manual mode (Fig. 5).

This paper is representing an HMI for the process of heating and quenching, Here JOG mode is used to manually operate the machine in case of any warnings, etc. RESET is to close all the functions and it makes the machine to restart from the beginning position, while in AUTO mode the machine will work automatically without the use of human hands (Fig. 4). Rotation 1 is representing rotation of heating Bed or Bed 1, So when it is rotating it will show positive else will show negative, In the same manner ROTATION 2 is representing Quenching Bed or Bed 2 while Top center Clamp/Unclamp represents clamping and unclamping of job so, when



Fig. 6 HMI showing clamping error

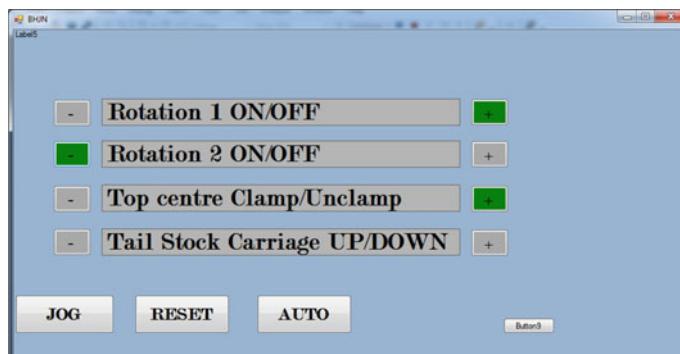


Fig. 7 HMI showing rotation of bed 1 or heating process

the job is at heating bed it will rotate and to give support clamper clamps the job. If the clamper grips the Job then it will show positive (Fig. 7) else with negative it will give a warning (Fig. 6).

When the system shows this warning, it will give us an idea that clamper is not holding a job so, to remove this error reset the processing and through JOG mode place the job in the right position. During the heating process, first clamper should clamp the job then only rotation 1 can happen, So during heating Rotation 1 and Top center Clamp/Unclamp will show positive.

During Quenching process, Rotation 2 will show positive (Fig. 8), and in this process, heated job is treated with water or chemical (coolant) to make it hard then before processing.



Fig. 8 Front view of a quenching process (left) and HMI showing the rotation of bed 2 or quenching process

3 Result and Future Scope

The implementation of HMI using VB.NET is successful for the process of hardening and quenching, now it can be implemented for further processes too. In automatic mode whatever the part of the machine doing the task this GUI is showing that task in real-time manner with the indication of positive and negative.

The problem solving is now become easier as we can solve errors by resetting and by JOG mode. For an indication of error an alarm is used so, the user can understand when it is facing an error. Now, in future we wish to implement SCADA using VB.NET so, the user can even have a look to whole plant through his chamber.

4 Conclusion

With this implementation using VB.NET, small-scale industries can use low budgets and efficient HMI system for their machines with little knowledge of working. As the competition among industries is increasing rapidly it will help small-scale industries to give high throughput in the small budget along with this it is simple to understand we can even visualize what process is going inside the chamber without harming us.

References

1. Implementation of customised SCADA for Cartoner packaging machine for cost effective solution. Int. Res. J. Eng. Technol. (IRJET)
2. Kumar, R., Kumar, A.: Design and hardware development of power window control mechanism using microcontroller. In: International Conference on Signal Processing and Communication, IEEE Xplorer, pp. 361–365
3. Yi, W., Yunfeng, Q., Jie, Y.: Application of automobile power window controller for peripheral interface controller. In: International Conference on Electronic Measurement & Instruments (ICEMI), August 2011, vol. 2, pp. 363–367. IEEE (2011)

4. Kumar, R., Ahuja, N.J., Saxena, M., Kumar, A.: Modelling and simulation of object detection in automotive power window. *Indian J. Sci. Technol. (IJST)* **9**(43) (2016). ISSN (Print): 0974-6846 ISSN (Online): 0974-5645
5. Design and implementation of hydraulic motor based elevator system. In: 6th IEEE India International Conference on Power Electronics (IICPE-2014), 8–10 December 2014, NIT Kurukshetra
6. Automated precise liquid transferring system. In: 6th IEEE India International Conference on Power Electronics (IICPE-2014), 8–10 December 2014, NIT Kurukshetra
7. Conceptual design of an automatic fluid level controller for aerospace applications. In: ICSNS 2015, 25–27th February 2015, Coimbatore, Tamil Nadu
8. Storage optimization of automated storage and retrieval systems using breadth first search algorithm. In: 1st International Conference on Nano-electronics, Circuits & Communication Systems (NCCS-2015), 9–10th May 2015, Ranchi, Jharkhand, India. Lecture Notes in Electrical Engineering, vol. 403. Springer, Berlin (2015)
9. White paper on Unlocking the power of .NET for industrial automation. International. www.ingeardrivers.com (2009)
10. www.plc-scada-dcs.blogspot.com
11. Nizamuddin, T., Iqbal, B.: Design and implementation of locally developed SCADA a cost effective solution. PCSIR, Institute of Industrial Electronics Engineering (IIEE), Karachi 16 (2010)

A Method to Reduce the Effect of Inter-Sub-band Interference (ISBI) and Inter Carrier Interference (ICI) in UFMC



Rajaraao Manda and R. Gowri

Abstract To support frequency segmentation and multi-service applications in fifth-generation wireless communications, the universal filtered multi-carrier (UFMC) modulation scheme is chosen and approved by worldwide mobile and wireless communication groups. However, the performance of UFMC-based systems mainly depends on the sub-band filter response, which influences the interference due to sub-bands and adjacent sub-carriers. The conventional UFMC uses the length of cyclic prefix/channel length as sub-band filter length, irrespective of the sub-band size. It may result in more number of computations to determine the filter response and hence introduces delay on the transmitter side. In addition, the filter tails extend to the one or many numbers of adjacent sub-bands in case of short sub-band sizes turns in performance degradation. In this paper, a novel method is proposed, which uses the straightforward concept of filter design to adapt the filter length based on the size of the sub-band. By this approach, the computational complexity might be reduced, Inter-Sub-band Interference (ISBI) and Inter Carrier Interference (ICI) can be reduced. Hence, the system gives better BER performance.

Keywords 5G · UFMC · ISBI · ICI · SNR · Filter length · Sub-band size · FIR filter

1 Introduction

The 5th-Generation (5G) cellular networks are envisioned to provide high data rates, superior QoS to end user, low latency (around 1 ms), and lower energy consumption.

R. Manda (✉)

Department of Electrical and Electronics Engineering, UPES, Dehradun, India
e-mail: rmand@ddn.upes.ac.in

R. Gowri

Department of Electronics Instrumentation & Control Engineering, University of Petroleum and Energy Studies, Dehradun, India
e-mail: rgowri@ddn.upes.ac.in

The other important application of 5G wireless systems will be the Internet of Things (IoT) and Device To Device (D2D), etc. [1]. To realize such a network, a number of innovative techniques involving at the level of various network layers to achieve the 5G targets. At the level of physical layer, the modulation along with its multiple access techniques defines the transmission performance of the network. For building up high data rate systems, flexible and scalable bands, multicarrier modulation is a very attractive technology. The most attractive multicarrier modulation scheme that executed in present broadband wireless systems is Orthogonal Frequency-Division Multiplexing (OFDM) by its flexible and scalable bands, and robustness in frequency selective channel environment. The OFDM seems inadequate in terms of new 5G aspects such as large side lobes; bandwidth utilization is limited by the length of Cyclic Prefix (CP), power amplifier complexity at RF system end due to large Peak to Average Power Ratio (PAPR), synchronization errors and for some applications OFDM cannot support like cognitive radio networks, IoT, etc. Therefore, the required 5G aspects have motivated the study of new modulation formats [2].

The evaluation toward next-generation waveforms, certain performance constraints required to be established. Such salient constraints are bandwidth flexibility for short packet transmission, spectral utilization, side lobe radiation, multi-user interface feasibility, backward compatibility, latency and complexity for implementation. In the past few years, a different type of modulation formats has been proposed for 5G air interface as an alternative to OFDM [3]. The Filtered Bank Multicarrier (FBMC) [4], Generalized Frequency-Division Multiplexing (GFDM) [5], Filtered OFDM (F-OFDM), Universal Filtered Multicarrier (UFMC) [6] are alternative modulations schemes have been proposed recently for the 5G air interface. The FBMC, which is the generalizations of OFDM. In order to overcome from large side lobes problem, each subcarrier is individually filtered using filter bank in FBMC. The FBMC provides low out of band radiation (side lobes) hence the inter carrier interference becomes smaller [7]. In spite of these advantages, the FBMC have more complexity in the design of the transmitter and receiver in practice. This is because of the order of filter required in FBMC is large and hence introduce an extra delay in addition to the hardware complexity. As per the 5G network requirements, it is infeasible for next-generation wireless communications in terms of short packet transmission.

In UFMC, a group of subcarriers is filtered (called sub-band filtering) to reduce the order of the filter [6] and it is also suitable for dynamic spectrum allocation in cognitive radio (CR) networks. Compared to the OFDM, UFMC provides lower side lobe radiation, achieve low latency and capable of supporting segmented frequency spectrum, also suitable for short packet transmission [8]. The F-OFDM is one of the 5G waveform candidatures to enable the 5G applications [9]. In F-OFDM, the sub-bands are filtered after the addition cyclic prefix (CP), where the size of Inverse Fast Fourier Transform (IFFT) and CP lengths are different for each sub-band and use the filter length is greater than or equal to the half of the symbol duration (i.e., half of the IFFT size) for good frequency localization. Which may result more complex at the transmitter/receiver side and the filter tails may extend to the next OFDM symbol. In UFMC, the filter length is short compared to the F-OFDM and equal to the CP length, results in less interference on the next symbol. The computational complexity is one of the significant drawbacks of UFMC by its filtering operation.

Recently, some methods are proposed to condense the computational complication of the transmitter/receiver in UFMC [10], which concentrated on the algorithms to reduce the number of multiplications and addition at the level of filtering and DFF/IDFT. The computational complexity can be reduced further by using these algorithms along with the possible reduction of filter length. The filter length (filter response ramp) affects the interference on the next symbol/sub-band, this value is either CP length or some numerical values those taken greater than the channel length [5]. In addition, the ramps/out band radiation introduced by filter response at both sides of the sub-band depends on the length of filter, which is more than a certain percentage of wireless channel delay spread to neglect the ICI and ISBI. The out band radiation can be reduced using some pulse shaping techniques [11], which may make further computational complexity. The method proposed in [12] uses separate subcarriers for interference cancellation and the weights of these subcarriers are adapting and optimizing to maximize the overall Signal-to-Interference Noise Ratio (SINR) under the power constraints. In [13], it is shown that the variation of inter sub-band interference for shorter and longer filter length with different sub-band intervals. Due to the use of sub-band interval, the spectral efficiency reduced. Here, we propose a novel approach to determine the order of filter based on the size of the sub-band using the basic design methodology of the FIR filter. The filter length is adapting to control ICI and ISBI without any sub-band interval.

The remaining part of the paper is as follows: Sect. 3 carries out the discussion of UFMC system model and Sect. 3 focuses on problem formulation in terms of ICI and ISBI, and includes the proposed method to adapt the filter length, Sect. 4 discusses the result analysis and performance comparison and lastly Sect. 5 concludes the proposed method.

2 System Modelling

2.1 UFMC Transmitter

The block diagram of the UFMC transmitter is illustrated in Fig. 1, where the effective number of data subcarriers (SC) N_{DC} are divided into ' N_{SB} ' number of sub-bands and each sub-band carries ' K ' number of subcarrier, i.e., $N_{SB} = N_{DC}/K$. After mapping the input QAM data samples to each sub-band, IFFT is performed for each sub-band individually. The p th sub-band QAM data symbols before IDFT block are defined as

$$X_p(k) = \begin{cases} X(Kp + k) & \text{for } k = 0, 1, \dots, K - 1 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where $0 \leq p \leq N_{SB}$ and $X(k)$ is the input QAM data sequence for $k = 0, 1, \dots, N_{DC}$

After N -point IDFT, the corresponding sub-band symbols are individually filtered by the filter of length L . The output of the p th sub-band filter can be defined as

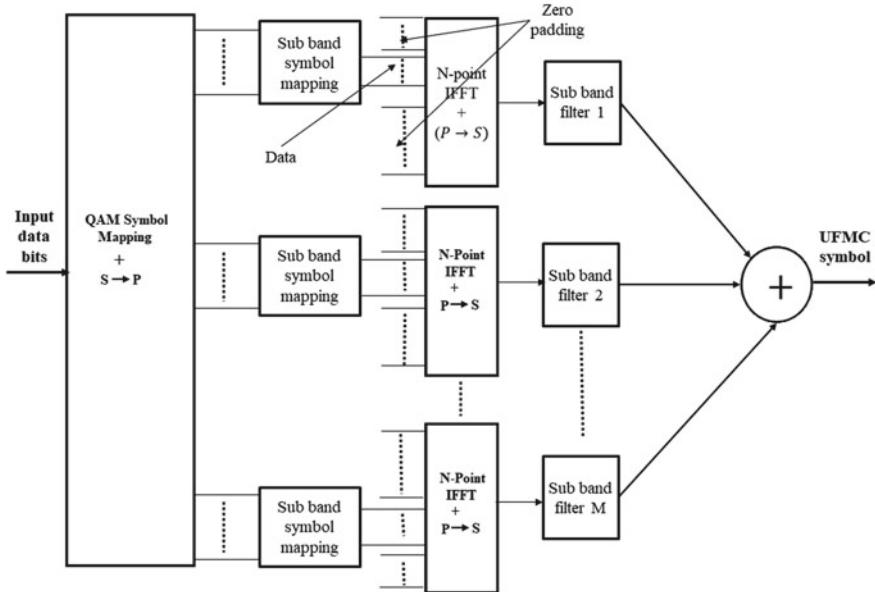


Fig. 1 Universal Filtered Multicarrier (UFMC) transmitter block diagram

$$y_p(n) = x_p(n) * f_p(n) \quad (2)$$

where $*$ is the convolution sum operator.

$$y_p(n) = \sum_{l=0}^{L-1} f_p(l)x_p(n-l) \quad (3)$$

$$\text{Where } n = 0, 1, 2, \dots, N + L - 2$$

where $x_p(n)$ and $f_p(n)$ are the N-point IDFT block of p th sub-band data symbols consists K number of subcarriers and filter impulse response of the p th sub-band filter (i.e., prototype filter response is shifted in frequency corresponding to the sub-band), respectively. $x_p(n)$ and $f_p(n)$ can be expressed as

$$x_p(n) = \frac{1}{N} \sum_{k=0}^{K-1} X_p(k) e^{j \frac{2\pi}{N} (Kp+k)n} ; n = 0, 1, \dots, N-1 \quad (4)$$

and

$$f_p(l) = h_d(l) e^{j \frac{2\pi}{N} (N_{vsc} + (p - \frac{1}{2})K + \frac{N}{2})l} \quad (5)$$

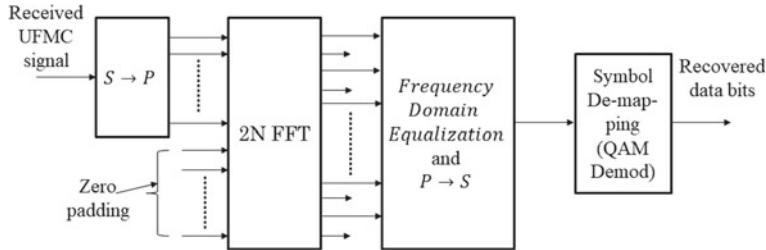


Fig. 2 Universal Filtered Multi-Carrier (UFMC) transmitter block diagram

where $p = 0, 1, 2, \dots, N_{SB}-1$ and $n = 0, 1, \dots, L-1$; N_{vsc} = number of virtual subcarriers for half band $[(N-N_{DC})/2]$ and $h_d(n)$ is the prototype filter impulse response. The final UFMC symbol expressed as the sum of the outputs of all sub-band filters.

$$y(n) = \sum_{p=0}^{N_{SB}-1} y_p(n) \quad (6)$$

The final UFMC symbol has a length of $N+L-1$, transmitted through the wireless channel.

2.2 UFMC Receiver

Figure 2 depicts a receiver block diagram of UFMC. At the receiver, the zeroes are padded to the received symbol and performed 2 N-point FFT then downsampled by 2 in order to estimate the received symbols on each subcarrier individually.

3 Problem Formulation and Proposed Method

3.1 Problem Formulation

Due to the filtering operation, the orthogonality between the subcarriers cannot be maintained at all. It makes the problem of interference between the adjacent subcarriers and sub-bands. The energy of the UFMC signal expressed as

$$E_{UFMC} = \sum_{n=0}^{N+L-2} |y(n)|^2 = \sum_{n=0}^{N+L-2} \left| \sum_{p=0}^{N_{SB}-1} y_p(n) \right|^2$$

$$= \sum_{n=0}^{N+L-2} \sum_{p=0}^{N_{SB}-1} y_p(n) y_p^*(n) + \sum_{n=0}^{N+L-2} \sum_{p=0}^{N_{SB}-1} \sum_{q=0}^{N_{SB}-1} y_p(n) y_q^*(n) \quad (7)$$

$q = 0$
 $q \neq p$

The first term in Eq. (8) represents the total sub-band energy for $p = q$ and the second term is inter-sub-band interference energy for $p \neq q$

$$E_{UFMC} = E_{SB} + E_{ISBI}$$

The sub-band energy of a particular sub-band E_{SB} can be divided into two components: one is the energy on the subcarriers and second is the interference on a particular subcarrier due to the other subcarriers known as ICI.

$$E_{SB} = E_{SC} + E_{ICI}$$

where

$$E_{SC} = \sum_{n=0}^{N+L-2} \sum_{p=0}^{N_{SB}-1} \sum_{k=0}^{K-1} \frac{1}{N^2} |X_p(k)|^2 \sum_{l=0}^{L-1} |f_p(l)|^2 \quad (8)$$

and

$$E_{ICI} = \sum_{n=0}^{N+L-2} \frac{1}{N^2} \sum_{p=0}^{N_{SB}-1} \sum_{k=0}^{K-1} \sum_{\substack{m=0 \\ m \neq k}}^{K-1} \frac{1}{N^2} X_p(k) X_p^*(m) E_{fp} \quad (9)$$

E_{ISBI} is the inter-sub-band interference can be expressed as

$$E_{ISBI} = \sum_{n=0}^{N+L-2} \sum_{p=0}^{N_{SB}-1} \sum_{q=0}^{N_{SB}-1} \sum_{\substack{k=0 \\ q \neq p}}^{K-1} \sum_{m=0}^{K-1} \frac{1}{N^2} X_p(k) X_q^*(m) E_{fpq} \quad (10)$$

$$\begin{aligned} E_{fpq} &= \sum_{l_1=0}^{L-1} \sum_{l_2=0}^{L-1} f_p(l_1) f_q^*(l_2) e^{j \frac{2\pi}{N} (Kp+k)(n-l_1)} e^{-j \frac{2\pi}{N} (Kq+m)(n-l_2)} \\ &= \sum_{l_1=0}^{L-1} \sum_{l_2=0}^{L-1} f_p(l) f_q^*(l) e^{j \frac{2\pi}{N} (Kp+k)(n-l)} e^{-j \frac{2\pi}{N} (Kq+m)(n-l)} \end{aligned}$$

Finally, the energy of the UFMC signal consists, the energy on subcarriers, the energy due to the inter carrier interference (ICI) and the energy due to the inter-

sub-band interference (ISBI). From Eqs. (9) and (10), it is clear that both E_{ICI} and E_{ISBI} depend on the filter response (also depends on filter length (L)) and the size of the sub-band K . The energy/power spectrum extends into the adjacent sidebands affected by the length of the filter (L) and which depends on the stopband attenuation of the filter. That means the filter length and stopband attenuation play a critical role to control the ISBI and ICI in the UFMC symbol. In case of large L and small K , the side lobe power may spread into an integer multiple of sidebands and hence increase the ISBI. This is clearly observed from the simulation results. To overcome this issue, we proposed a method which explained in next sections along with the simulation results.

3.2 Filter Design

The filter length in UFMC is defined such that the delay spread of channel should be less than a certain amount of percentage and this length is fixed for any value of sub-band size K . However, the filter length influences the ISBI and ICI from Eq. (9) and (10), respectively. For a smaller size of sub-band K the spreading or out of band radiation prolonged into a number of sub-bands and hence ISBI. For $K > L$, this extension will be less. If L is varying according to the value of K , it may possible to maintain ISBI within a range and it can be insignificant value.

Here, we proposed a method to design filter response with an adaptive filter length based on the sub-band size. The FIR filter can be designed by windowing methods and the length of the filter can be determined by formulae shown in Table 1 [14]. The length of the filter mainly depends on the stopband attenuation and the normalized transition width (Δf_N).

To determine the filter length, consider the UFMC signal shown in Fig. 3, having N number of subcarriers and six sub-bands. The UFMC symbol is repeated after symbol duration. The transition width (Δf) of the filter is the difference between the stopband frequency (f_s) and passband frequency (f_p).

According to the basic filter design principle, the stopband frequency is less than or equal to the half of the sampling frequency. From Fig. 2, it can be expressed as

Table 1 Different type of windowing methods and their specification

S. No	Type of windowing	Stopband attenuation A (in dB)	Filter order/length
1	Hanning	41	$N = 3.1/\Delta f_N$
2	Hamming	53	$N = 3.3/\Delta f_N$
3	Blackman	74	$N = 5.5/\Delta f_N$
4	Kaiser	>21	$N = \frac{A-8}{14.36\Delta f_N}$

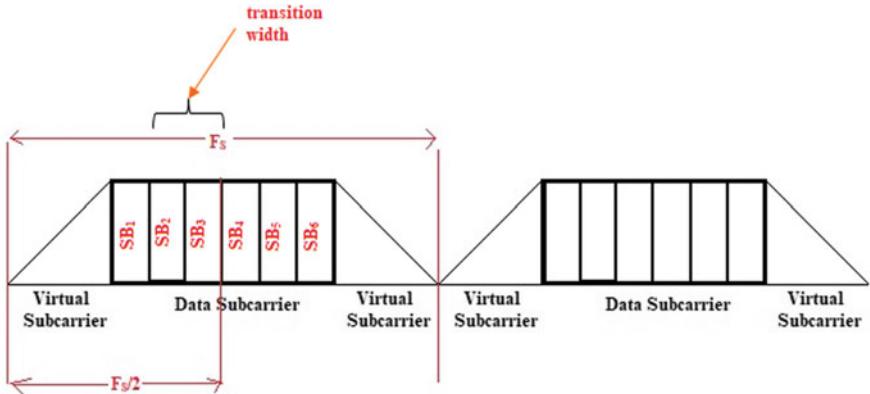


Fig. 3 UFMC symbol divided into six sub-bands

$$\text{transition width } \Delta f = f_s - f_p = \frac{F_s}{2} - BW_{SB} - BW_{VSC} \quad (11)$$

where BW_{SB} and BW_{VSC} are the bandwidth assigned for sub-band, it is defined as $BW_{SB} = Kf_{sc}$ and bandwidth consumed by virtual subcarriers and defined as $BW_{VSC} = N_{VSC}f_{sc}$, respectively; N_{VSC} represents the number of virtual subcarriers. The normalized transition width $\Delta f_N = \Delta f/F_s$. The filter length is calculated by best widow method from Table 1. Or the following formula:

$$L = \frac{-10 \log(\delta_p \delta_s) - 13}{14.36 \Delta f_N} \quad (12)$$

where δ_p and δ_s are the passband and stopband ripples, respectively. Here we consider both are equal. The filter coefficients are calculated by Dolph–Chebyshev window method [15].

4 Simulation Results

Here we consider physical resource blocks (PRB) that use data transmission in LTE, LTE-A, and 5G: New Radio (NR), which consists of 12 (SC). For 3GPP: NR, the subcarrier spacing (SCS) can be defined by $2^\mu \times 15$ kHz, where $\mu = 0, 1, 2, 3, 4$. As per the reference [16] the bandwidth configuration for 5G is presented in Table 2 for 15 kHz SCS.

The sub-band length in UFMC considered the multiple of PRB for transmission all types of data services. One or two PRBs used for short burst transmission and more number of PRBs can be used for long bursts transmission. In conventional UFMC-based systems [6, 8], the filter length is equal to the length of the channel

Table 2 Bandwidth configuration for 15 kHz SCS

SCS kHz	Bandwidth/specifications	5 MHz	10 MHz	20 MHz	50 MHz	100 MHz
15	Number of subcarriers	333	666	1333	3333	NA
	PRBs/data subcarrier	25/300	52/624	106/1272	270/3240	NA
	FFT/IFFT size	512	1024	2048	4096	NA
	Sampling rate (MHz)	7.68	15.36	30.72	61.44	NA
	CP duration/length	36	72	144	288	NA

Table 3 Simulation parameters

Parameter	Value
Channel bandwidth	10 MHz
Modulation type	64-QAM
FFT size	1024
CP length	72
Number of symbols	14
Sub-band size	12, 24, 48 and 60
Stopband/passband ripple factor	0.001
Channel type	AWGN

or the length of the cyclic prefix plus one. This filter length is fixed for any size of the sub-band. Nevertheless, in this paper, the filter length is altered for different sub-band sizes as shown in Table 3.

The difference in power spectral density (PSD) for the proposed and conventional method can be clearly observed in Figs. 3 and 4, that's how the PSD extended into other sub-bands and the variation of sidelobe attenuation. When the sub-band size is 12 SC and $L = 73$ (CP length +1), the Out of Band Emission (OBE) of sideband is extended to many numbers of sidebands. In case of adaptive filter method ($L = 23$), the OBE is more but limited to some number of sidebands. The simulation results produced for 10 MHz channel bandwidth with sub-band size 12 and 60, 64-QAM are depicted in Figs. 4 and 5.

There are 52 sub-bands if each sub-band carries 12 SC (i.e. one PRB). The total cumulative ISBI is more in case of fixed filter length compared to the adaptive filter length scheme. It observed from the simulation result shown in Fig. 6a.

The energy of the ICI signal in UFMC is plotted and depicted in Fig. 6b, where the ICI increases with increase in sub-band size. Here it is seen that, for fixed filter length, the ISBI is overriding ICI (because ICI is an insignificant value) when the sub-band size is small. For long sub-band size, the ICI is dominated the ISBI, it can be observed from simulations result depicted in Fig. 6. Therefore, the SNR versus BER performance degraded in case of longer sub-band size compared to the small size of sub-band. The simulation results are depicted in Fig. 7b for parameters mentioned in Table 3. But in the case of the proposed algorithm, the change in ISBI and ICI is very

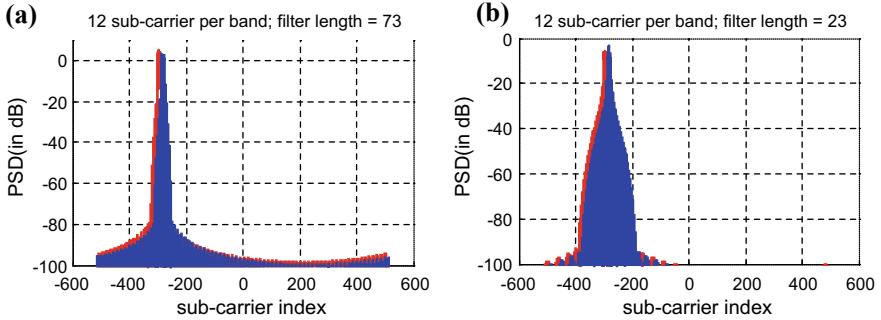


Fig. 4 **a** and **b** Power spectral density of two adjacent sidebands of UFMC signal with fixed filter length $L = 73$ and proposed adaptive filter length ($L = 23$) respectively; sub-band size = 12 SC

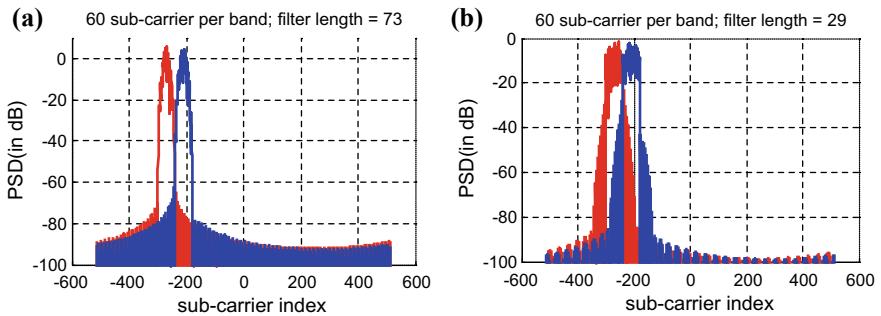


Fig. 5 **a** and **b** Power spectral density of two adjacent sidebands of UFMC signal with fixed filter length $L = 73$ and proposed adaptive filter length, respectively; sub-band size = 60 SC

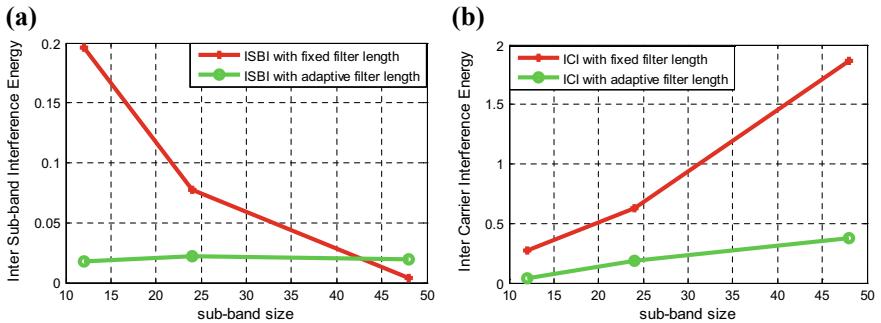


Fig. 6 **a** and **b** Average ISBI per sub-band in UFMC symbol and average ICI energy per sub-band, respectively

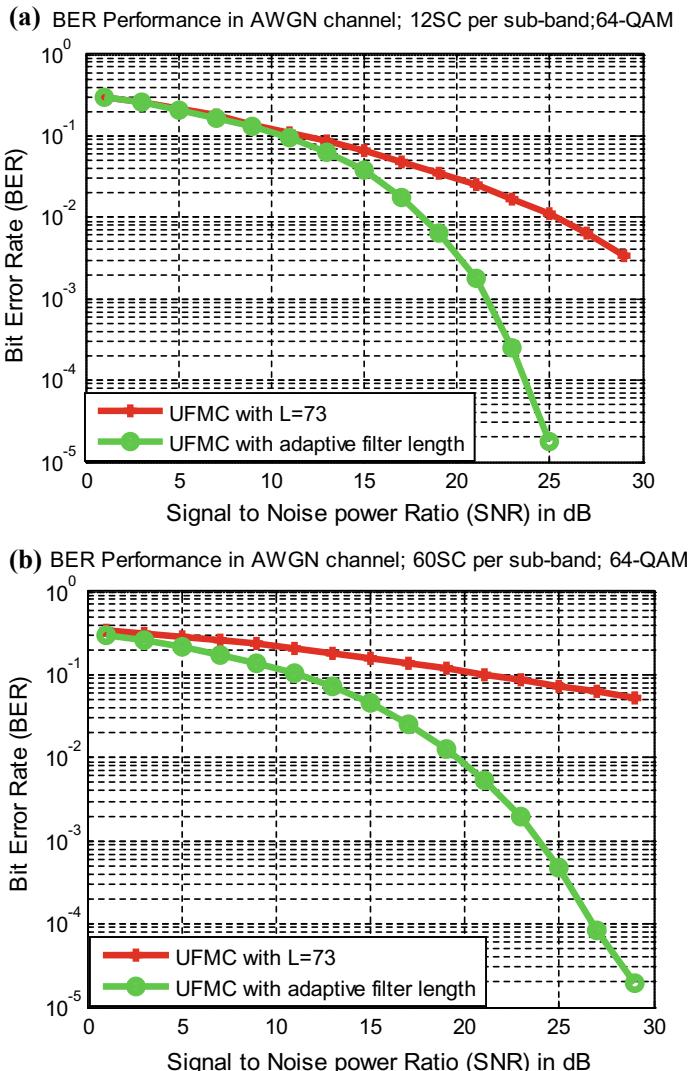


Fig. 7 a and b UFMC based system BER performance for sub-band size = 12 SC and sub-band size = 60 SC respectively

small and almost constant. For acceptable BER 10^{-2} , the minimum required SNR is approximately 25 dB for fixed filter length algorithm and 18 dB for the proposed adaptive length algorithm. Therefore, we can save approximately 7 dB compared to the conventional method.

5 Conclusion

A non-orthogonal multicarrier transmission is well suited for 5G networks to support especially for multi-user interface services such as D2D, machine-type communication, IoT, etc. The UFMC is one of the best-suited methods for this. However, the filter response influences the orthogonality between the subcarriers and interference between the adjacent sub-bands. In this paper, the interference was controlled and minimized by adapting the filter length with respect to the size of sub-band without using any interval between the sub-bands. Also, the variation of ISBI and ICI with respect to sub-band size and filter length is analyzed. The simulation results proved that the ISBI and ICI are less and almost constant with respect to the sub-band size compared to the conventional method. In addition, the SNR and BER performance is good in this scenario.

References

1. 3GPP: Requirements for further advancements for evolved universal terrestrial radio access (E-UTRA) (LTE-Advanced), vol. 9, pp. 1–17 (2010)
2. Banelli, P., Buzzi, S., Colavolpe, G., Modenini, A., Rusek, F., Ugolini, A.: 5G networks: who will, pp. 80–93 (2014)
3. Wunder, G., et al.: New Waveforms for New Services in 5G, 1st edn. Elsevier Ltd, Amsterdam (2017)
4. Farhang-Boroujeny, B.: Filter bank multicarrier modulation: a waveform candidate for 5G and beyond. *Adv. Electr. Eng. Adv. Electr. Eng.* **2014**, 1–25 (2014)
5. Michailow, N., Datta, R., Krone, S., Lentmaier, M., Fettweis, G.: Generalized frequency division multiplexing for 5th generation cellular networks. *IEEE Trans. Commun.* **62**(9), 3045–3060 (2014)
6. Vakilian, V., Wild, T., Schaich, F., Ten Brink, S., Frigon, J.: Universal filtered multi-carrier technique for wireless systems beyond LTE. In: IEEE Globecom Work, pp. 223–228 (2013)
7. Farhang-boroujeny, B.: OFDM versus filter bank multicarrier. *IEEE Signal Process. Mag.* **28**, 92–112 (2011)
8. Schaich, F., Wild, T., Chen, Y.: Waveform contenders for 5G - suitability for short packet and low latency transmissions. In: IEEE 79th Vehicular Technology Conference (VTC Spring), pp. 1–5 (2014)
9. Zhang, L., Ijaz, A., Xiao, P., Molu, M., Tafazolli, R.: Filtered OFDM systems, algorithms and performance analysis for 5G and beyond. *IEEE Trans. Commun.* **66**(3), 1205–1218 (2017)
10. Nadal, J., Abdel Nour, C., Baghdadi, A.: Novel UF-OFDM transmitter: significant complexity reduction without signal approximation. *IEEE Trans. Veh. Technol.* 1–1 (2017)
11. Mukherjee, M., Shu, L., Kumar, V., Kumar, P., Matam, R.: Reduced out-of-band radiation-based filter optimization for UFMC systems in 5G. In: 2015 International Wireless Communications and Mobile Computing Conference (IWCMC), pp. 1150–1155 (2015)
12. Zhang, Z., Wang, H., Yu, G., Zhang, Y., Wang, X.: Universal filtered multi-carrier transmission with adaptive active interference cancellation. *IEEE Trans. Commun.* **65**(6), 2554–2567 (2017)
13. Duan, S., Yu, X., Wang, R.: Performance analysis on filter parameters and sub-bands distribution of universal filtered multi-carrier. *Wirel. Pers. Commun.* **95**(3), 2359–2375 (2017)
14. Schafer, R.W., Oppenheim, A.V.: Discrete-Time Signal Processing. PHI Private Limited (1996)

15. Lynch, P.: The Dolph-Chebyshev window: a simple optimal filter. *Mon. Weather Rev.* **125**(4), 655–660 (1997)
16. 3GPP TS 38.104: 3GPP technical specification group radio access network: NR base station (BS) radio transmission and reception, vol. V15.1.0, No. Release 15, pp. 1–133 (2018)

Real-Time Coordination with Electromagnetic Relay and Advanced Numerical Relays in Laboratory Environment/Setup



Isaac Ramalla, Ram Mohan Sharma and M. A. Ahmed

Abstract In this paper, smart grid trends like technical, socio-economic development to the society are emphasized which can implement the fact that dynamic interaction and related infrastructure development are designed to overcome the challenges in power system. This paper summarizes on a state-of-the-art power lab in which distribution network setup is created and all normal system parameters are implemented and analyzed. It creates a distribution automation model and explains the importance of smart substations and its application in distribution systems.

Keywords Smart substation road map · Grid of the future · Self-healing distribution system · Canadian electricity association

1 Introduction

In this paper, a brief outline of several key aspects of a smart substation is discussed. These include the Communication Infrastructure of a Smart City, the Role of Relays and Switchgear in a Smart Substation, and preventing failures through Smart Substations. In addition to this, the Power Systems Lab at a university was related to smart substations as a functional smart substation model. Few experiments were performed which highlighted the role of advanced relay protection systems in a smart substation. Cybersecurity failures and error due to failure are also discussed at the end of this paper.

I. Ramalla (✉)
MLRITM, Hyderabad, India
e-mail: hodeee@mlritm.ac.in

R. M. Sharma
UPES, Dehradun, India

M. A. Ahmed
ACE, Aurangabad, India

2 Communication Infrastructure of a Smart Substation

With the development of smart substations and smart cities, a robust and efficient communication infrastructure must also be developed to support many of the new requirements that come with it. The main requirement of a smart substation communication infrastructure is that it must support two-way communication between the supplier and consumer. An effective method of data and energy transfer (as shown in Fig. 1) and proper relay protection systems must also be implemented so that data on faults and errors within the substations are reliably communicated in a timely manner to resolve the issue due to the faults.

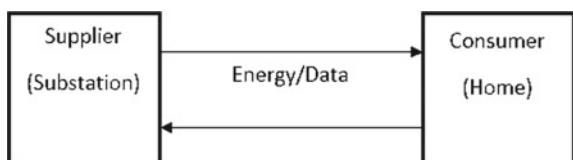
2.1 Two-Way Communication

The concept of two-way communication in the application of a smart substation is a key element of its functionality. The ability to have supplier–consumer transfer of both energy and data is what makes the infrastructure “smart”. In this paper, we are discussing data transfer between two substations. We are concentrating on how the substations can improve, control, and make the substation more reliable [1]. In this section, smart metering and fault identification will be discussed as they relate to the two-way communication aspect of a smart substation communication infrastructure.

2.2 Smart Metering

With the development of smart substations, there is a need for a more intelligent and efficient metering infrastructure. The concept of smart metering in a smart substation communication infrastructure refers to real-time demand and usage which acts as feedback from the consumers in the substation to the suppliers; data from a smart metering system will be sent to the substation control center where it can be analyzed to both make real-time adjustments as well as predict future consumption trends. This allows the supplier to optimize energy production to reduce waste and increase efficiency. With smart metering implemented properly, consumers could effectively monitor their energy usage, allowing them to reduce the cost for their own benefit as well as reduce waste and pollution from excess energy being gener-

Fig. 1 Smart substation model



ated. With the development of smart metering, there is a potential for the integration of smart appliances in homes, which would interact with the two-way communication network. Smart appliances could provide even more data such as the source of wasted energy, and be controlled to prevent energy shortage to the supplier present at consumer end by data energy transfer concept (shown in Fig. 1).

2.3 *Substation Layout*

In traditional power substations, the layout is typically centralized. Both power generation and control stations are centralized, all power is generated at one location and distributed to the rest of the substation, all data is received and processed at one location. In a smart substation, there is a significant shift from this traditional layout. With the increase in relays and remote terminal units (RTU), identification of faults [1] in a real-time substation at a central hub is complex. Data transfer capabilities will also be needed across the substation, each relay is monitored and controlled between substations [2]. Huge data transfer needs to be monitored without any trials and errors.

2.4 *Electric Power Systems Lab as Smart Substation Model*

Advanced power system laboratory layout of the UPES Electric Power Systems Lab is shown in Fig. 2. In this lab, a smart substation is emulated. The numerical relay systems are identified as the protection system in the substation, and the electromagnetic relays are acting as Fault Passage Indicators (FPI). This model shows us how a smart substation would function using these technologies, with much experimentation, this will provide useful information as it relates to the implementation of a real smart substation. With this setup, we can analyze various fault response scenarios as they would occur in real life and prepare viable response techniques in a risk-free environment, making a future transition to a smart substation more streamlined and reliable. As a fault is injected into the relays, we will see the real-time effects through EnerVista (SCADA) software, and be able to model the response techniques using these experiments (Fig. 3).

3 **Role of Relays and Switchgear in Smart Substation**

As smart substations, advance relays become implemented around the world, everyone must understand how they work. A smart substation is a substation-like structure that provides advanced substation reliability analysis, failure protection, security, privacy protection services and allows one section of the power substation to be shut

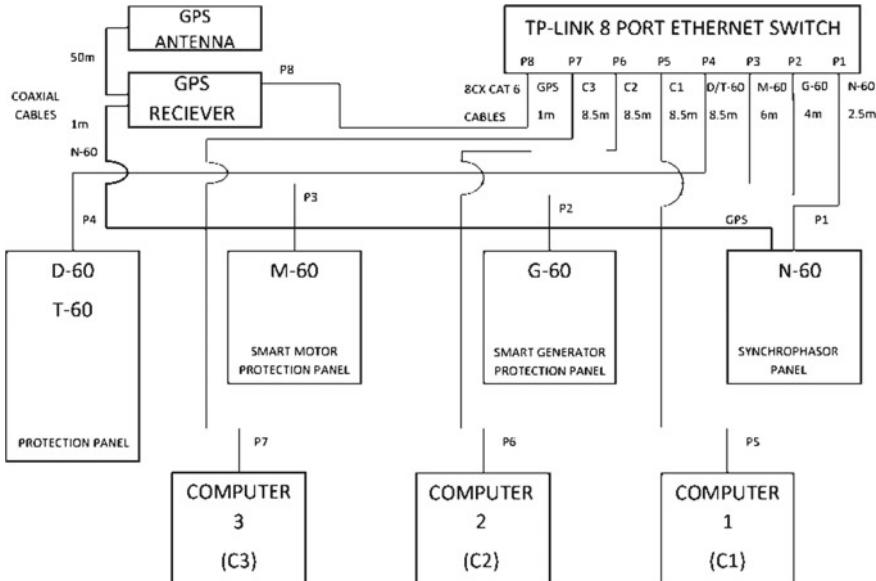
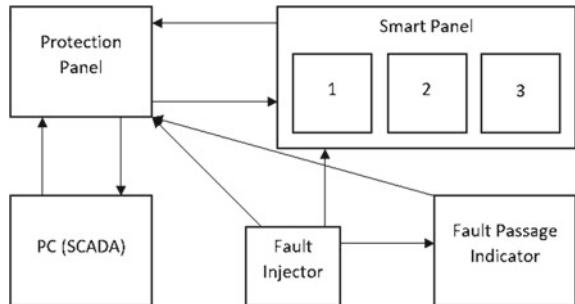


Fig. 2 Architecture of Power System Lab

Fig. 3 Power System Lab smart grid model



off or the power be redirected if needed. This is usually done when there is either a failure or damage to the area. This allows us to prevent cascading failures or damages, if not isolate them altogether. When the smart substation notices a failure or damage, a protection relay displays an alarm or error message, allowing a worker to see and get help to fix the alarm or event failure as fast as possible. The relay records the failure in a digital folder, noting the date, time, location, and what type of failure was, allowing access whenever needed, and keeping track of year-round failures, and when a fault is cleared it records the fault clearing time, and displays the fault report. This allows as a fault report history for someone to work in a substation. It records when the worker arrived and fixed the problem at hand. For some failures, it is possible for some smart substations to fix themselves by using smart protection units (Numerical Relays). For example, if there was a failure in one section of the

substation, the ideal smart substation setup could get the data required to determine there was a failure, then either shut off power to that section through the switchgear, change the path of the power supply, to allow nearby devices to still get power but not where the failure is, or automatically trigger help, allowing workers to solve the issue at hand in less period of outage time.

3.1 Department Power Systems Lab Setup

In the Power Systems Lab at UPES, we have simulated a smart substation system which has an integrated protection panel, three smart panels (Smart Motor Protection Panel M60, Smart Generator Protection Panel G60, and a Synchrophasor Panel N60), all of which is connected to a 220-volt DC power source. We then have eight electromagnetic relays (Single Line to Ground Fault, Over Voltage Relay, Thermal Relay, I.D.M.T. Over Current Relay, 3 Phase Line to Line Fault, Performance of Long Transmission Line Under Different Loads—Ferranti Effect, Percentage Differential Relay, and an Earth Fault Relay) which we use as Fault Passage Indicators by injecting fault currents and voltages through the secondary injector kit. It allows to understand the relay response and analyze the fault conditions. The integrated protection panel and the three relays are all connected to a router and GPS allowing us to access their stored data and control them from a remote location.

3.2 Communications in Smart Substation Simulation

In the UPES Power Systems lab, there is the architecture for a simulated smart substation. There are several numerical and electromagnetic relay systems which will represent the substation and individual cities, representing smart substation model. When a fault is passed through one of the electromagnetic relays, a signal is sent to the integrated protection panel that a fault has been detected. This information is then sent to the SCADA system and is recorded as a fault event. LED indicators are configured to show where the fault has been detected. Once the fault is removed, the relay can self reset and the data is recorded in the SCADA software.

As part of the first analysis on relay response, a fault was created in the “Panel for Single Line to Ground Fault”, which will be represented as “Toronto”. As the fault was created, a signal was sent to the “Integrated Protection Panel”, which is communicated with the SCADA PC. One of the LED indicators which were configured to Toronto now displays that there is a fault. In a real-life scenario, a team would be sent to resolve the issue at the Toronto substation. Once this fault has been removed manually by the authorized field personal, the SCADA main console operator at the Ontario Control Station resets the relay, identifying that the system is in healthy condition and the color of the relay and circuit breaker is change to normal status.

4 Injecting Fault Currents and Voltages

4.1 Experimental Work and Procedure

For us to simulate a real-time failure at a smart substation, we injected a fault current into the three-phase line to line fault electromagnetic relay panel, which acts as a fault passage indicator to the integrated protection panel. Line to Ground Fault electromagnetic relay was set to trigger when the current varies by 0.4 amps.

Case-I (Open Test): We passed 110 volts through an open circuit, which resulted in ‘no-fault’ being detected due to less current being passed through.

Case-II (Short-Circuit Test): We then short-circuited the transformer secondary side, which is acting as the line to line fault on the overhead line.

Case-III (Relay by pass-biasing): Applying 3.5 volts resulted in 1.85 amps which would damage the equipment connected to the power lines, however, this was conducted without the relay system engaged.

Case-IV (Relay under Operation): With the relay on, we applied 1 V which resulted in 0.38 amps, which did not trip the fault, we then applied 1.5 volts which resulted in 0.48 amps, since it was higher than 0.4 amps (relay pickup current), which triggered the fault. The electromagnetic relay then sent a fault message through the relay contact signal source to the integrated protection panel allowing the contact input (H2C – On Status) to be triggered as shown in Fig. 4 and the red color is highlighted to the operator (an indication of alert).

This simulates a real-time example by the Single Line to Ground Fault relay resembling a power system in a city. A system in Toronto is shown in Fig. 5, in which when the system detects a failure, it sends a fault message to integrated protection panel in a control room which could locate anywhere, and the system either does the required steps as discussed in above four cases to fix the failure automatically or alerts a worker or operator in the control room to get the help needed to fix the issue by the field operator at Toronto. All the data is recorded at that particular instant and transferred to all substations about the fault, so the connected substations take necessary actions to heal and address the fault.

5 Preventing Failures Through Smart Substations

5.1 Failures Due to Cybersecurity Parameters

Smart substations can provide numerous benefits when implemented correctly, such as reduced energy waste and improved efficiency. One of the main challenges with smart substations is the cybersecurity issue, with a more advanced communication infrastructure, with many more intelligent devices spread across the substation, such

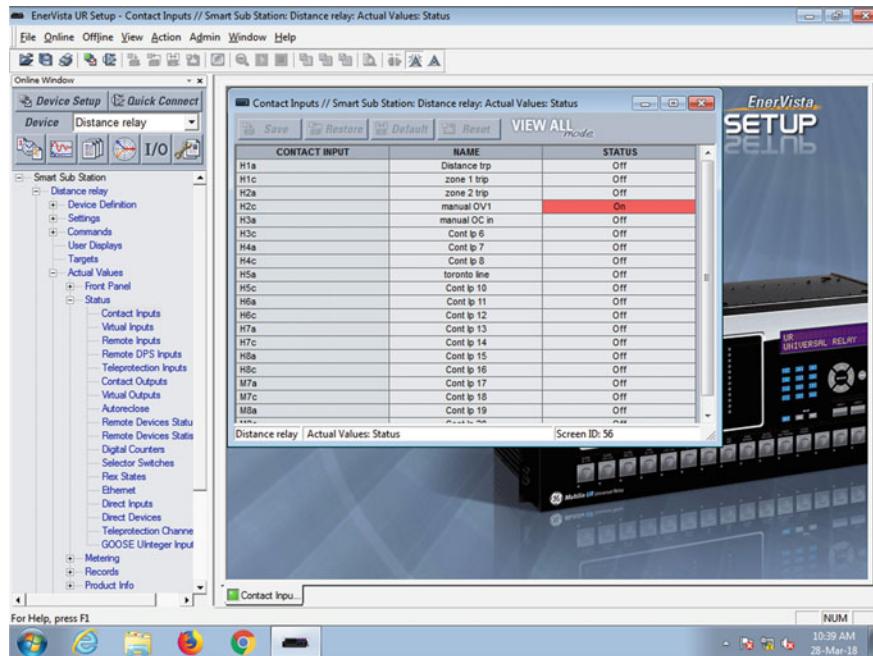


Fig. 4 Control input status of numerical relay

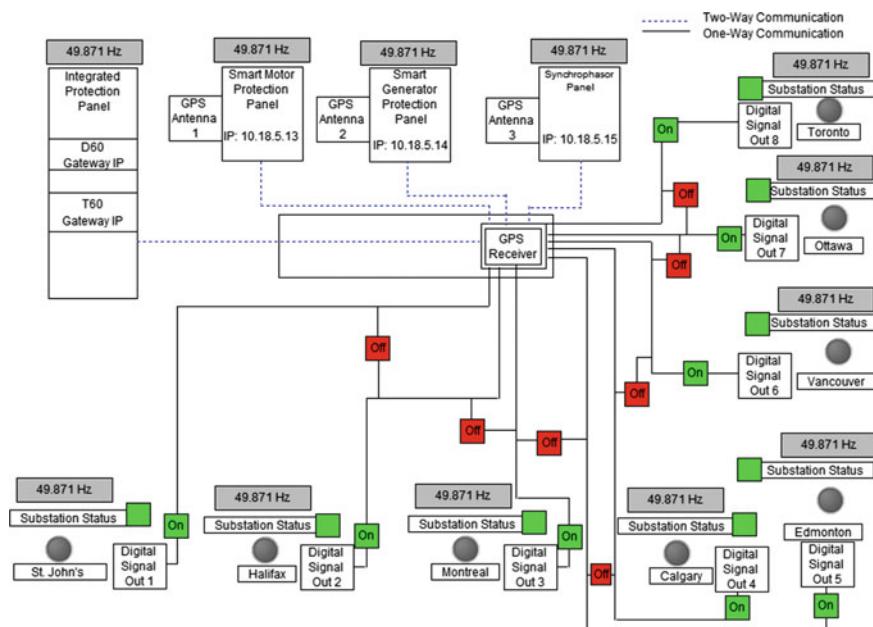


Fig. 5 Smart grid system in Toronto grid lab layout on EnerVista SCADA software (operator view)

as smart metering systems and relay protection systems, there is vastly increased potential for cybersecurity threats and attacks. In order for a smart substation to be properly implemented, there must be a robust and efficient cybersecurity system in place. Some of these potential attacks could be in the form of a False Data Injection attack (FDI), where false data is injected into the communication network in order to decrease the integrity of the data collection and relay protection systems, as well as a denial of service attack which basically is the flooding of the communication network with data in order to prevent system malfunction temporarily [3]. In the event of a network interruption, whether intentional or not, the status of the relay system will change. This can be observed onsite or from a remote location through SCADA software. Wrong interpretation of the fault is detected by the system and leads to malfunctioning of switchgear during healthy conditions and making the system unreliable and wastage of energy.

5.2 Failures due to Electrical Parameters

The smart substations are effective and reliable there are many complications that can cause failures. A smart protection system must be able to protect the smart substation from user errors, equipment failures, natural disasters, and cyber-attacks [4]. It does this by predicting weak points in the substation and preventing failures before they have the chance to occur, or if a failure has already occurred, the smart protection system will identify, diagnose, and recover from the failure if possible, thus making the smart substation self-healing. All of this is done as quickly as possible, so no cascading failures occur due to a previous failure. One of the most common failures is a path failure, which is when something goes wrong with the delivery of power to the target, for example, if the wires break and there is a complete loss of power. This is one example where a smart protection system will be able to fix itself. The system will recognize the failure, automatically alert someone, and change the path of power flow if possible. The ideal design for a smart substation is to divide the power substation into smaller, autonomous microsubstations, allowing for the substation to function normally, however if a failure occurs, you can cut off a much smaller portion of the substation than you normally can, if these substations are designed correctly, cascading events or further system failure can be avoided altogether.

5.3 Other Failures

There are many other causes of failures, such as recording failures, which occur when there is an error when recording the data from smart meters. Also, failures can be caused by using old or damaged equipment/appliances. Sometimes if a protection system is set to be too sensitive, false failures can occur, this is when the relay thinks something is wrong, so it records the failure and does what is needed to solve the

issue at hand, when nothing is wrong. Another major issue is the hidden failures. A hidden failure occurs when the incorrect system states or control actions are triggered by another system event. These are considered one of the main sources of large-scale disturbances, for example, these may cause large cascading failures such as the Northeast blackout in 2003, which was initially caused by a false relay trip.

6 Conclusion

In this paper, we discussed the concepts and applications of a smart substation. We discussed the smart substation working in contrast with normal substations and self-healing techniques followed without human intervention and reliable and quick decisions are taken through smart substations. The authors have mimicked smart substation as in real time at Power Systems Lab to show the readers, researcher and students to understand several real-life scenarios followed in normal and smart substations during healthy (normal) and faulty (emergency) situations. We can also analyze how faults would be handled and solved using a modern smart-infrastructure and the authors also discussed how this advancement can lead to unwanted failures due to misuse of sensing devices. In future, researchers can work on IEEE C37.117 standards and make the system more reliable and self-secure.

Acknowledgements The authors acknowledge Department of Electrical and Electronics of University of Petroleum, Dehradun, India to conduct the experiment in power System Laboratory and we also acknowledge department laboratory staff for laying out the cables as per the design layout, we thank work-term students from Memorial University from Canada for helping us in coding and decoding the electrical and network cabling and documenting them in the relay software. We acknowledge the Lab Head, Department Head and Electrical Department Faculty Members from University of Petroleum, Dehradun, India for helping/guiding us to perform the experiments on the advanced relay modules successfully.

References

1. Al-Kandari, A., Gilany, M., Madouh, J.: An accurate technique for locating faults by distance relays. *Electr. Power Energy Syst.* 477–484 (2018)
2. Chen, C.R., Lee, C.H., Chang, C.J.: Optimal over current relay coordination in power distribution system using. *Electr. Power Energy Syst.* 217–222 (2018)
3. Al-Kandari, A., Gilany, M., Madouh, J.: An accurate technique for locating faults by distance relays. *Electr. Power Energy Syst.* 477–484 (2018)
4. Kowalik, R., Rasolomampionona, D.D., Januszewski, M.: Laboratory testing of process bus equipment and protection functions. *Electr. Power Energy Syst.* 405–414 (2018)

Symmetric Cryptography and Hardware Chip Implementation on FPGA



Priyanshi Vishnoi, S. L. Shimi and Adesh Kumar

Abstract Cryptography is the process of creating the data secretly to provide secure communication in the existence of third parties called adversaries. Cryptography changes data into a format that is not readable for an illegal user. The paper provides the hardware chip-based solution for cryptographic encryption and decryption using the XOR cipher technique. The chip design is done in Xilinx ISE 14.2 with the help of VHDL language and simulation is carried for 8-bit to 256-bit encryption/decryption in Modelsim 10.0 software. The FPGA parameters such as hardware and timing are also synthesized for SPARTAN 3E, 3s500efg320-5 device.

Keywords Cryptographic encryption and decryption · FPGA synthesis · XOR cipher technique · VHDL language

1 Introduction

Cryptography [1, 2] provides secure information and communication techniques based on mathematical concepts and several algorithms to change the data in ways that are very hard to crack. The main features of cryptography are confidentiality, integrity, nonrepudiation, and authentication. Confidentiality refers that actual person only can understand the information. Integrity refers that the information cannot be changed. Nonrepudiation refers that the transmitter cannot deny his/her intentions in the communication of the information at a future stage. Authentication is the main feature, which means that transmitter and the receiver should confirm each other. Cryptographic-based approach can be a good solution for security in the commu-

P. Vishnoi · S. L. Shimi

Department of Electrical Engineering, National Institute of Technical Teachers Training and Research (NITTTR), Chandigarh, India

A. Kumar (✉)

Department of Electrical and Electronics Engineering, School of Engineering, University of Petroleum and Energy Studies, Dehradun, India

e-mail: adeshmanav@gmail.com

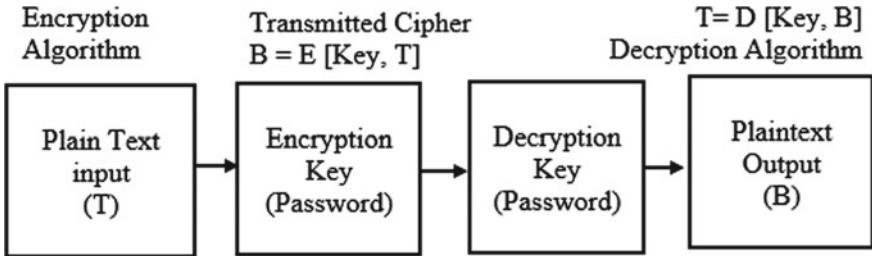


Fig. 1 Cryptographic encryption/decryption

nication system to achieve protected information and secured communication. In the communication network, there are electronics devices embedded with lost cost hardware. They can follow the cryptographic approach of encryption and decryption with public-key cryptography. The design of encryption and decryption scheme is essential to provide data integrity and data confidentiality in the smart grid. In Cryptography, encryption the actual message or data is referred to as plain text which is encoded with encryption key, called ciphertext [3] and communicated over a channel. Description is the contrary process, in which the ciphertext is decoded to get the actual plain text. With the help of decryption key and ciphertext, the decryption technique provides the original plain text.

The model of cryptography is shown in Fig. 1 in which plaintext (T) is encrypted with key value (Key) and transmitted ciphertext is $B = E [key, T]$, the same text is extracted with decryption.

2 Cryptography Techniques

Cryptography [4, 5] is the techniques that involve encryption and decryption with the sharing of the same key at both ends or the different key on both ends. There are mainly two types of encryption algorithms [6] called symmetric and asymmetric algorithm. Symmetric key algorithm [8] is also called a private key algorithm and the symmetric key algorithm is called public-key algorithm. In a private key algorithm, there exists only one key for both encryption and decryption algorithm. The complexity of the private key algorithm is less and easier to implement for higher speed applications and can be implemented with hardware. In asymmetric cryptography approach, both encryption and decryption processes use a different key and difficult to implement having a complex structure.

The example of symmetric cryptography and asymmetric cryptography is shown in Figs. 2 and 3. The user only knows the key. In the symmetric key approach, the encryption key and encryption key both are the same but in asymmetric key, the encryption and encryption keys both are different. In Fig. 2 symmetric key ABCD@#123 is used on both encryption and decryption end whereas Fig. 3 is

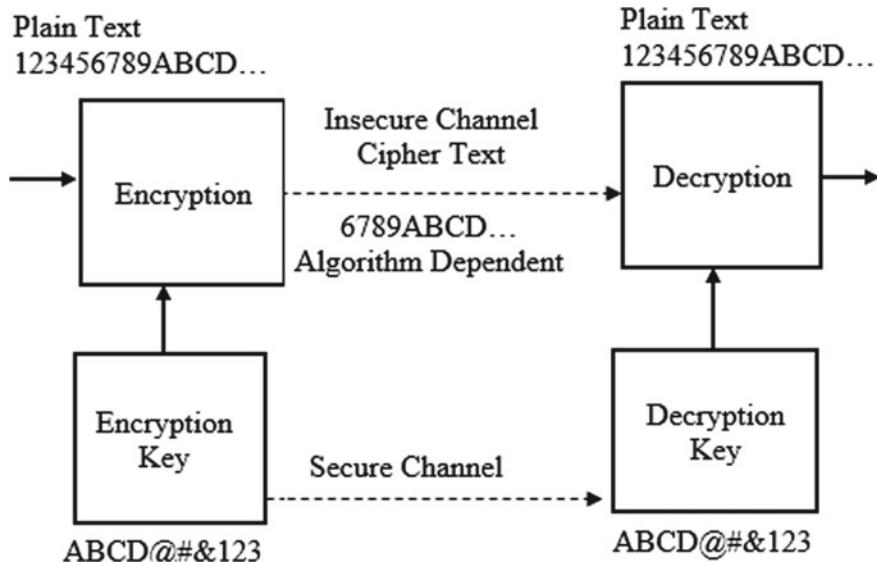


Fig. 2 Symmetric Key encryption/decryption

following asymmetric key as DCBA@#123 and ABC@#1234 on encryption and decryption ends respectively. The key size is a very important aspect to secure the data, long key size means the data is more secured. Encryption processes are very significant for cryptography with an approach of key management subsequently there are diverse algorithms that offer different levels of security based on the key size.

3 XOR Encryption and Decryption (XED) Method

In cryptography, the XOR encryption [7] is an additive cipher that can be applied on ‘n’ bit data encryption and decryption with ‘n’ bit key value. The algorithm is based on XOR gate operation or some times referred as modulus-2 addition.

$$0 \oplus 0 = 0$$

$$0 \oplus 1 = 1$$

$$1 \oplus 0 = 1$$

$$1 \oplus 1 = 0$$

$$A \oplus 0 = A$$

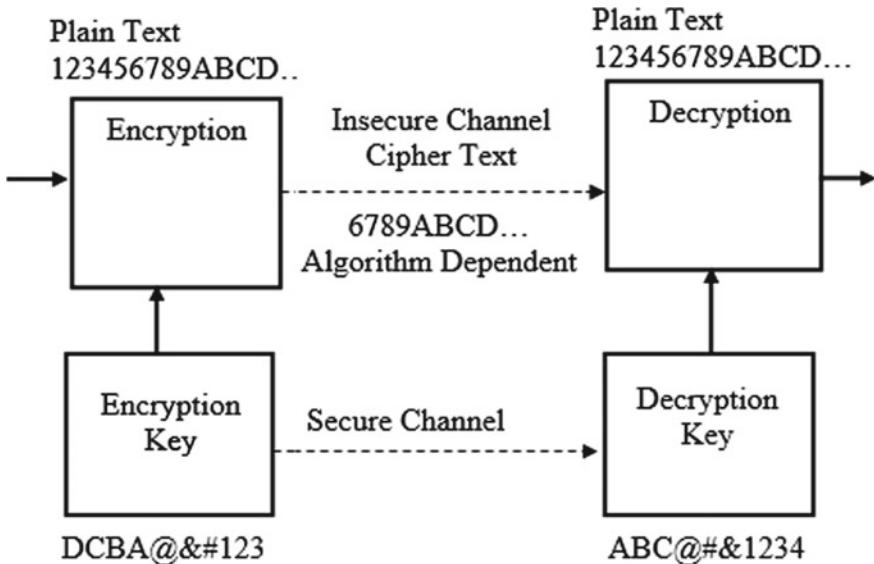


Fig. 3 Asymmetric Key encryption/decryption

$$A \oplus A = 0$$

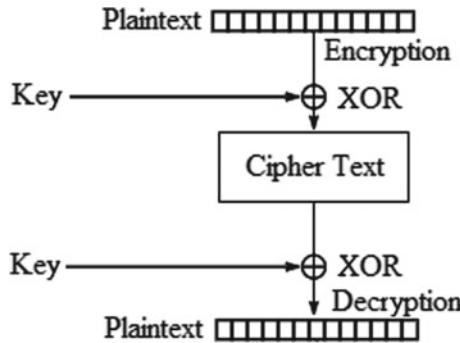
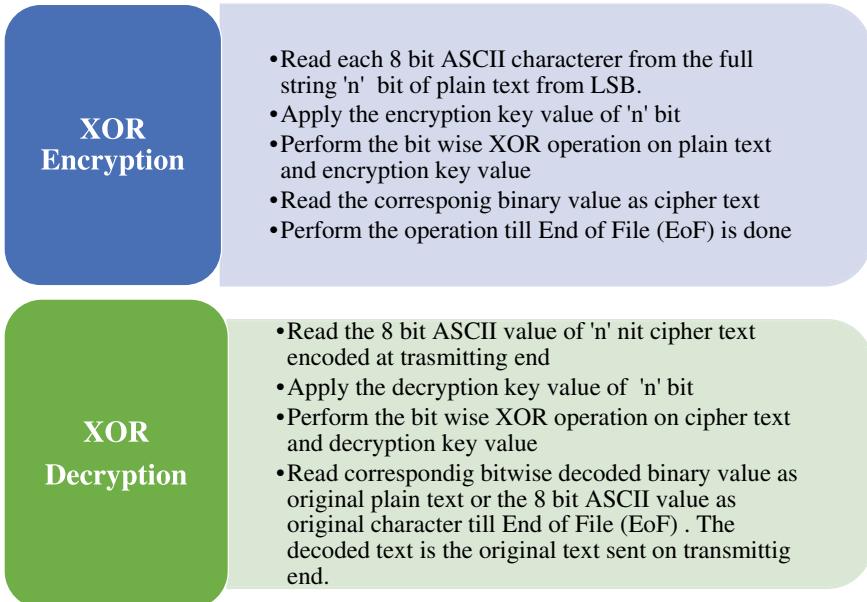
$$(A \oplus B) \oplus C = A \oplus (B \oplus C)$$

$$(A \oplus B) \oplus B = A \oplus 0 = A$$

The string of the data is given as the plain text and bitwise XOR operation is performed to encrypt the text at transmitting end to get the ciphertext [9, 10]. The XOR operation is further applied to the ciphertext with the same key and same plain text is decoded on receiving end. The processes depicted in Fig. 4 and steps are given in Fig. 5.

4 Results and Discussions

The hardware chip of XOR encryption/decryption is implemented in Xilinx ISE 14.2 software. RTL view of the design is shown in Fig. 6, which has all inputs and outputs used in the design. The detail of the pins is given in Table 1.

**Fig. 4** Bitwise XOR encryption/decryption**Fig. 5** Steps for 'n' bit XOR encryption/decryption

The modelism simulation waveform for 8 bit (1 byte), 16 bit (2 bytes), 32 bit (4 bytes), 64 bit (8 bytes), 128 bit (16 bytes), and 256 (32 bytes) bit encryption and decryption are shown in Fig. 7a–e. The simulation tests details is given in Table 2.

The chip is synthesized on the SPARTAN 3E 3s500efg320-5 device. Tables 3 and 4 presents the hardware and timing parameters summary for the developed design.

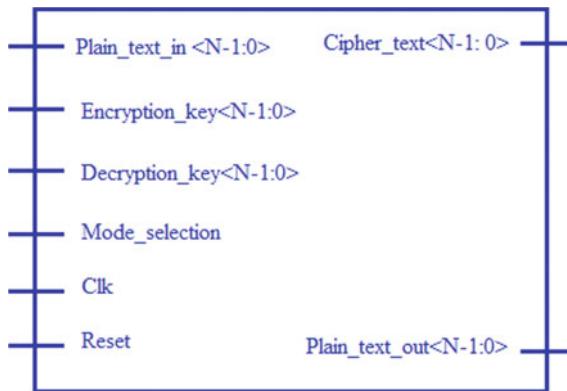


Fig. 6 RTL view of the developed chip

Table 1 RTL pin detail of designed chip

Pins	Direction	Description
Plain_text_in <N-1:0>	Input	It presents the input data as the textual text of the encryption end. The size of the text can vary in 'N' bit. In the simulation is considered of 8 bit to or 256 bit
Encryption_key <N-1:0>	Input	It presents the input encryption key as the password of the encryption end. The size of the text can vary in 'N' bit, considered of 8 bit to 256 bit
Decryption_key <N-1:0>	Input	It presents the input decryption key as the password of the decryption end. The size of the text can vary of 'N' bit, considered of 8 bits to 256 bits
Clk (1 bit)	Input	It is the input of 1 bit used to provide rising clock pulse to work digital logic at 50% duty cycle
Reset (1 bit)	Input	Reset pin is used to reset the logic circuitry and synchronized with clock pulse
Mode_selection (1 bit)	Input	It is 1-bit input in chip to decide chip operation in encryption mode only or in encryption/decryption. When mode selection = '0', chip performs encryption only. When Mode selection = '1', chip performs encryption/decryption both
Cipher_text <N-1:0>	Input	It is text, which is achieved at transmitting end after XOR encryption. It also varies from 8 bit to 256 bit
Plain_text_out <N-1:0>	Output	It presents the output data as the textual text of the decryption end. The size of the text can vary in 'N' bit. In our simulation, it is 8 bit to 256 bit

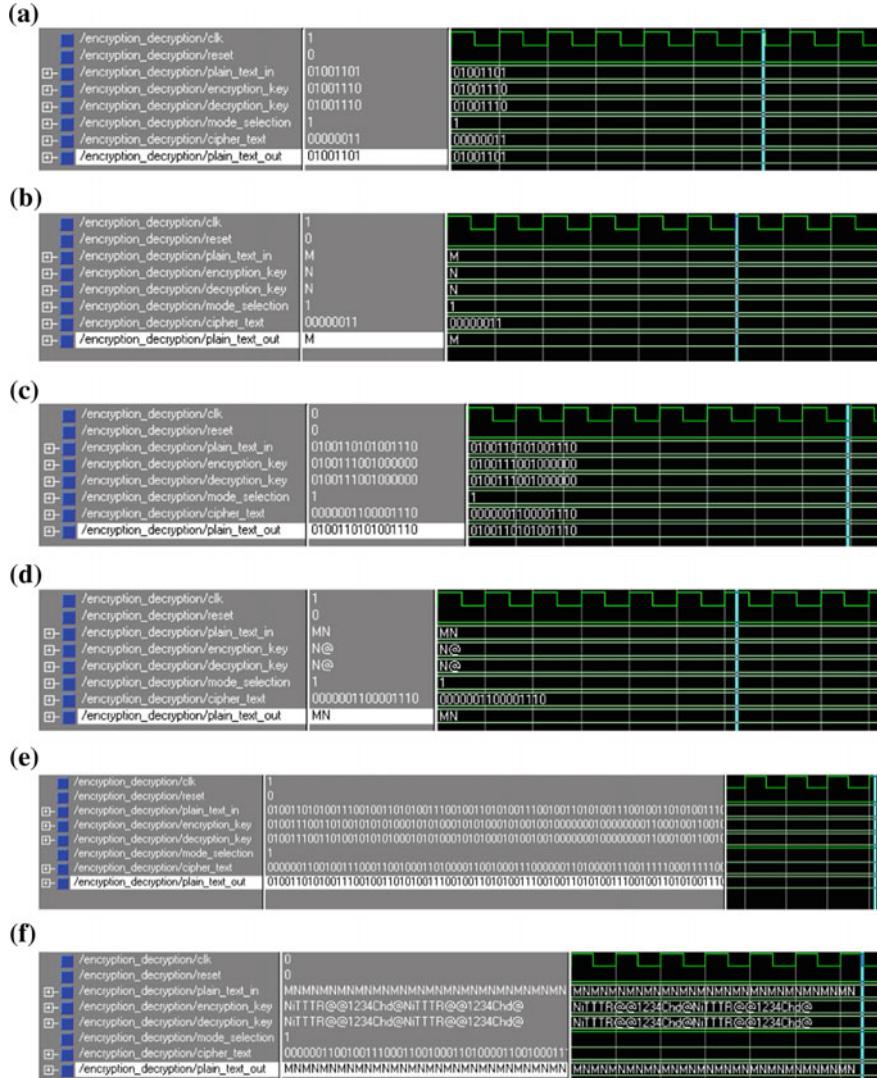


Fig. 7 Simulation waveform for encryption/decryption **a** 8 bits in binary **b** 8 bit in ASCII **c** 16 bits in binary **d** 16 bits in ASCII **e** 256 bits on binary **f** 256 bit in ASCII

Table 2 Simulation test cases

Test	Values
Case-1 (8 bit)	Clk = Clk, Reset = '0' Plain_text_in = "01001101" in binary, 'M' in ASCII Encryption_key = "01001110" in binary, 'N' in ASCII Mode_selection = '0' for encryption and Mode_selection = '1' for decryption Cipher_text = "00000011" in binary Decryption_key = "01001110" in binary, 'N' in ASCII Plain_text_out = "01001101" in binary, 'M' in ASCII
Case-2 (16 bit)	Clk = Clk, Reset = '0' Plain_text_in = "01001101 01001110" in binary, 'MN' in ASCII Encryption_key = "01001110 01000000" in binary, 'N@' in ASCII Mode_selection = '0' for encryption and Mode_selection = '1' for decryption Cipher_text = "00000011" in binary Decryption_key = "01001110" in binary, 'N@' in ASCII Plain_text_out = "01001101" in binary, 'MN' in ASCII

(continued)

Table 2 (continued)

Table 3 Device utilization summary for 256-bit encryption/decryption logic

Parameters	Utilization (256 bit)
Number of Slices	412 out of 4656 8.84%
Number of Slice Flip Flops	255 out of 9312 2.73%
Number of 4 input LUTs	810 out of 9312 8.7%
No. of bounded I/OBs	43 out of 232 18.53%
Number of GCLKs	1 out of 24 4.16%
Total memory usage (kB)	151290 kB

Table 4 Timing summary for 256-bit encryption and decryption logic

Timing parameter	Encryption (256 bit)
Frequency (Max)	275.00 MHz
Period (Min)	2.156 ns
Time before clk (Min)	5.726 ns
Time after clock (Max)	3.197 ns
Speed grade	-5

5 Conclusions

Cryptography is associated with the methods of converting normal plain text into unintelligible text and vice versa. Cryptography is not only protecting the data from hackers or alteration, but also applied for user authentication. The VHDL code is used to developed the XOR cipher based encryption and decryption chip hardware and synthesized on SPARTAN 3E FPGA for 1 byte to 32 bytes (8 bits to 256 bits) data encryption/decryption. The synthesized module is helpful for several types of security and imaging processing applications in which security is the primary concern.

References

1. Bellare, M., Boldyreva, A., Micali, S.: Public-key encryption in a multi-user setting: security proofs and improvements. In: International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Berlin, Heidelberg (2000)
2. Boldyreva, A. et al.: Order-preserving symmetric encryption. In: Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Berlin, Heidelberg (2009)
3. Churchhouse, R., Churchhouse, R.F.: Codes and ciphers: Julius Caesar, the Enigma, and the Internet. Cambridge University Press (2002)
4. Chelton, William N., Benissa, Mohammed: Fast elliptic curve cryptography on FPGA. IEEE Trans. Very Large Scale Integr. VLSI Syst. **16**(2), 198–205 (2008)
5. Kumar, Adesh, Kuchhal, Piyush, Singhal, Sonal: Secured Network on Chip (NoC) architecture and routing with modified TACIT cryptographic technique. Proc. Comput. Sci. **48**, 158–165 (2015)

6. Lindell, Y., Jonathan, K.: Introduction to Modern Cryptography. Chapman and Hall/CRC (2014)
7. Nag, A., et al.: Image encryption using affine transform and XOR operation. In: 2011 International Conference on Signal Processing, Communication, Computing and Networking Technologies. IEEE (2011)
8. Stalling, W.: Cryptography and Network Security, 4/E. Pearson Education India (2006)
9. Yalla, P., Jens-Peter K.: Lightweight cryptography for FPGAs. In: International Conference on Reconfigurable Computing and FPGAs, 2009. ReConFig'09. IEEE (2009)
10. Zodpe, H., Ashok, S.: An Efficient AES Implementation Using FPGA with Enhanced Security (2018)

Design of Microstrip Antenna with Extended Ground



N. Prasanthi kumari, P. Suresh Kumar, G. M. G. Madhuri
and B. Praveen Kitti

Abstract In this work, a novel microstrip antenna with the patch is compared with the patch with the ground extended of the same dimensions. The antenna with the extended ground has shifted frequency to higher side with better antenna performance is observed. A ground plane patches with 2.5 mm are placed on two sides of the patch. With this newer design, the frequency of the patch that is working at 3.6 GHz has shifted to 8.1–11.2 GHz (with better gain). This present antenna structure is fabricated using the lithographic technique. The proposed antenna provides a new way to improve physical ground plane more than the size of the original PCB ground unlike adding separate wires.

Keywords Microstrip patch antenna · Ground plane · Extended ground

1 Introduction

There is an increasing demand of smaller and lighter handset devices by the users, on the other hand, they want long battery life of the handsets. Antennas play a major part in the field of communications. They are used to transmit and receive the data without using wires. As the technology is advancing, the motive has been reduced to increase the efficiency and gain [1] of the antennas. Unfortunately, when the size of the handset is reduced and compactness is achieved, these slim and shortened devices do not have a good antenna performance. A ground plane is a flat horizontal conducting surface. A ground plane in an antenna reflects the radio waves from the other antenna elements. A ground plane on a printed circuit board (PCB) is a layer of copper foil. This ground plane is connected to the ground point of the

N. Prasanthi kumari (✉) · P. Suresh Kumar
Department of Electronics, Department of Mechanical, University of Petroleum and Energy
Studies, Dehradun, India
e-mail: prasanti@ddn.upes.ac.in

G. M. G. Madhuri · B. Praveen Kitti
Department of ECE, PSCMR College of Engineering, Vijayawada, Andra Pradesh, India

circuit, with the help of one terminal of power supply. In this project, we propose an approach of using a wire to electrically extend portions of the ground plane to improve the performance of the antenna. This project aims to design a microstrip antenna with the extended ground to achieve improved antenna performance without physically enlarging the ground plane (performance is improved in terms of gain and efficiency). Thus, the different parameters of the antenna will be known. The testing of the antenna will follow this. Hence, the desired antenna shall be obtained, thereby fulfilling the objective of the design of the system.

The substrate used in this design is FR4, which has a dielectric constant of 4.4. Probe feed technique is used to provide the required excitation and to achieve the desired results. The dimensions of the patch are calculated using transmission line equations. Here we will be using a patch of square shape and the following equation is used to calculate the dimensions of the patch which is 2.45 cm and the strip size is taken as 1 cm.

$$w = \frac{c}{f_o \sqrt{\frac{(\epsilon_r+1)}{2}}}$$

The designing part of the antenna would be done using 3D EM tools like HFSS. The designed antenna is then to be patterned over a double-sided printed circuit board and then etching is to be done in order to remove the extra part. The main advantage of these types of antenna is their low weight and low in cost. These antennas are economical to manufacture and the size of such antennas are small ever since they are used in high frequencies and as the band of the frequency or rather the operating frequency increases there is a reduction in the dimensions of the patch of the antenna. The gain of such kind of antennas [2] can range from 6 to 9 dB.

2 Design

A Microstrip Patch Antenna consists of a very thin metallic conducting plane over a substrate and in turn which has ground plane beneath the substrate. A dielectric substance separates the patch and the ground. The conductor normally used for the patch is copper. The patch can have any shape but for this present work, a square patch which is relatively simpler to design and the analysis and performance of the prototype. The substrate used is generally nonmagnetic. The relative permittivity of the substrate is taken into the key parameter because relative permittivity enhances the fringing field which rules radiation. The length L, width W, and thickness H of the antenna are shown in Fig. 1.

The chosen substrate is FR4 with a thickness of 1.762 mm and $\epsilon_r = 4.4$. The feed position is ideally placed at the center, while the ground plane below the printed circuit. The optimized geometry of the proposed antenna is fed by 50Ω SMA connector which occupies an area of $50 \times 50 \text{ mm}^2$ and is shown in Fig. 2.

Fig. 1 Rectangular Microstrip Patch Antenna

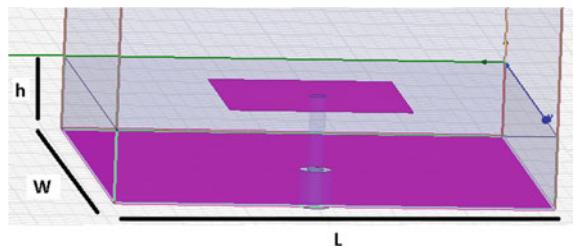
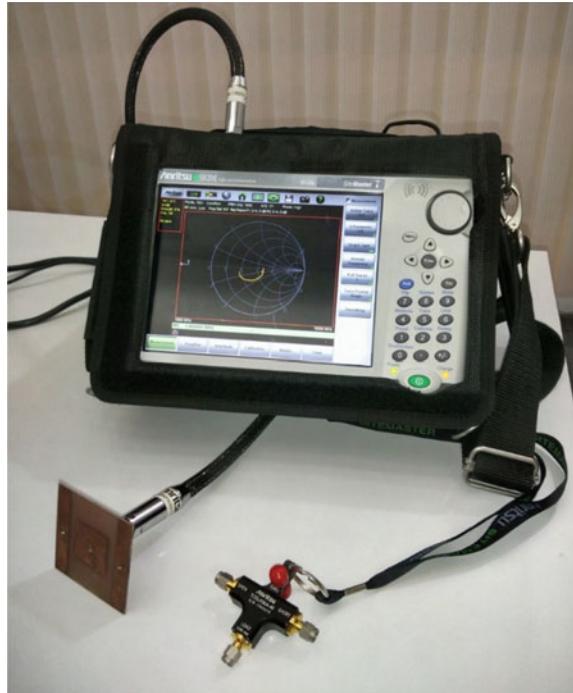


Fig. 2 Fabricated Patch Antenna with extended ground



This is the same size patch is taken with as is used in, which only achieved a starting frequency of 3.6 GHz. The measured and simulated results of starting frequency are shown in Figs. 3 and 4.

Without increasing the antenna size entire UWB range can be covered by the extended ground plane concept [3, 4]. The antenna used in [5] is used as a reference antenna to validate the proposed technique. The extended ground affords a further current path to enhance the low-frequency bandwidth [6]. By considering altered ground plane dimensions, most extreme transient currents flow along the edges of the infinite ground plane with impedance coordinating at 8.25 GHz resonance frequency. The measured and simulated results of the extended ground plane are shown in Figs. 4 and 7. For a ground plane of size $50 \text{ mm} \times 50 \text{ mm}$, there is an effect on return loss but bandwidth [7, 8] is the same and for bigger ground planes there will be a change

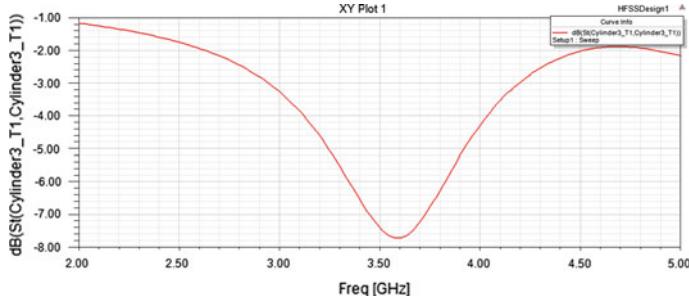


Fig. 3 S-Parameter for normal Patch Antenna designed in HFSS

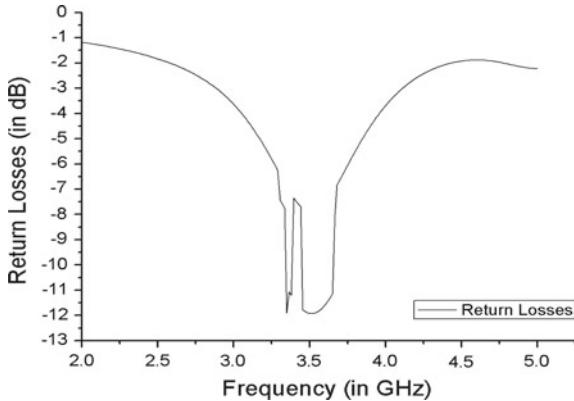


Fig. 4 Return Losses for normal Patch Antenna—Result

in impedance matching, so patch dimensions and feed location has to be changed. To correct the effect the extra plane is created on the lower part of the patch antenna with 1 mm conducting patches which is shown in Fig. 5. The upper part and lower part of the ground planes are connected by vias.

3 Results

The results of the normal patch antenna and antenna with the extended ground are presented in Figs. 3, 4 and 6, the latter is having better gain and resonating at higher frequency with 4 GHz shift towards the right. Figure 6 is a simulated result for the antenna with the extended ground and, Figs. 7 and 8 explain about the tested results of S_{11} and VSWR of the microstrip patch antenna, which is resonating from 8.1 GHz to 11.2 GHz.

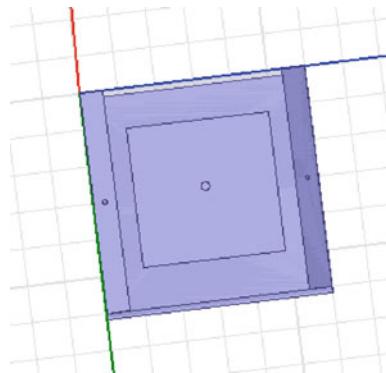


Fig. 5 MPA with extended ground

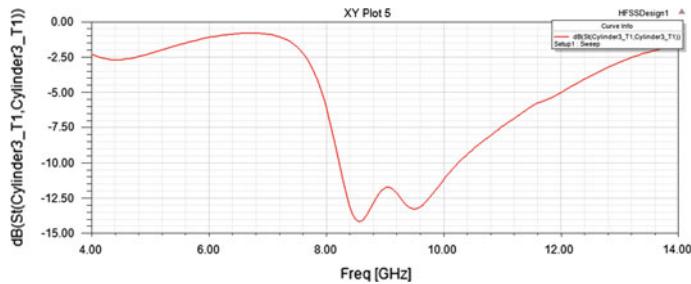


Fig. 6 S-Parameter for Patch Antenna with extended ground designed in HFSS

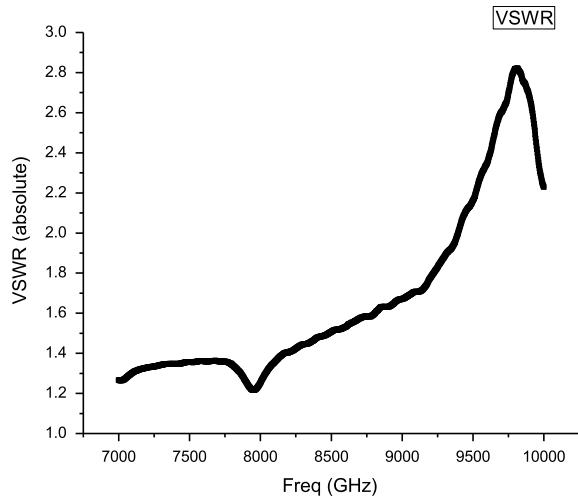
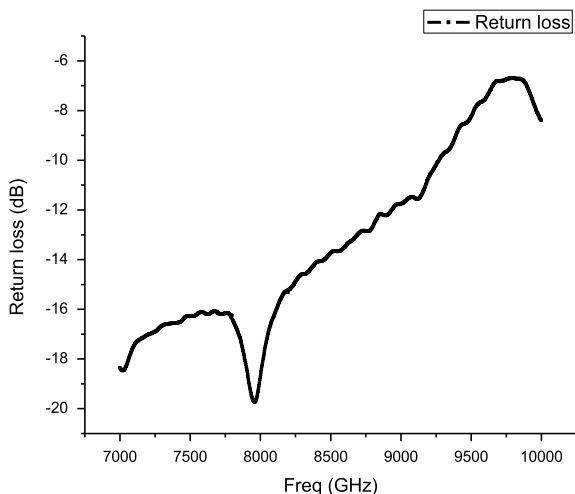


Fig. 7 VSWR for Microstrip Patch Antenna with extended ground designed in HFSS

Fig. 8 S_{11} for Microstrip Patch Antenna with extended ground Tested result



4 Conclusion

The outcome of compact size ground plane on MPA has been examined to broaden the antenna ground with no outer ground plane. This subject has been returned to in the point of view of utilization in portable communications and Wireless Local Area Network stations. For each and every ground size, the patch size is changed and for compact sized ground planes bandwidth and gain is affected. The microstrip antenna and the antenna with extended ground were constructed and functioned at the chosen frequency and power levels. The square patch antenna with the extended ground was designed and fabricated and the desired results were obtained. The obtained hardware and software results are the same as predicted results. Improved performance of the antenna in terms of gain was achieved by extending the ground. Also, there was a shift in frequency about 3.5 GHz and thus the antenna now can be operated around 8 GHz. This frequency lies in the X-band and hence can be used for radar applications and in phased arrays.

References

1. Mishra, R.G., Mishra, R., Kuchhal, P., Kumari, N.P.: Analysis of the microstrip patch antenna designed using genetic algorithm based optimization for wide-band applications. *Int. J. Pure Appl. Math.* **118**(11). Special Issue on Engineering and Applied Mathematics (2018). ISSN 1314-3395
2. Mishra, R.G., Mishra, R., Kuchhal, P., Kumari, N.P.: Optimization and analysis of high gain wideband microstrip patch antenna using genetic algorithm, *Int. J. Eng. Technol.* **7**(1.5) ISSN 2227-524X

3. Urban, R., Peixeiro, C.: Ground Plane Size effects on a Microstrip Patch Antenna for small Handsets, pp. 521–524. IEEE Conference 2004
4. Custódio, P.: Small ground plane size effects on microstrip patch antenna revisited. In: IEEE Conference 9–15 June 2007, Honolulu, HI, USA
5. Mishra, R., Mishra, R., Kuchhal, P., Kumari, P.: Performance Enhancement of Rectangular Microstrip Antenna by Inserting Notches and Slits, pp. 391–396 (2018). https://doi.org/10.1007/978-981-10-7122-5_39
6. Wang et al.: A Decoupling Technique Based on Partially Extended Ground Plane for Compact Two-Port Printed Monopole Antenna Arrays, GeMic March 10–12, 2014, Aachen, Germany, IEEE Conference 2014
7. Mishra, R.G., Mishra, R., Kuchhal, P., Kumari, N.P.: Design and analysis of CPW-Fed microstrip patch antennas for wide band applications. In: Proceedings of the IEEE International Conference on Inventive Computing and Informatics (ICICI 2017), Coimbatore, ISBN: 978-1-5386-4031-9, 23–24 November 2017
8. Mishra, R.G., Mishra, R., Kuchhal, P., Kumari, N.P.: Analysis of bandwidth and directivity of rectangular microstrip antenna by inserting slots, notches and slits. In: Proceedings of the 2nd IEEE International Conference on Communication and Electronics Systems (ICCES 2017), Coimbatore, ISBN 978-1-5090-5013-0, 19–20 October 2017

A Dual-Band High-Gain Patch Antenna with Reduced Cross-Polarization Levels Designed for X Band and K Band



Tanisha Gupta, Tarun Kumar and Rajeev Kamal

Abstract A novel Swastika-shaped printed antenna with suitable for X and K Band applications with higher gain than previously designed antennas was designed. Resonating at two center frequencies of 9.7 and 12.7 GHz, it has a bandwidth of 450 MHz around each resonant frequency. The width of four slots in the patch, the feed position have been optimized to increase the gain. The shortening of the ground plane resulted in a significant reduction in cross-polarization levels ultimately achieving high peak gain of 5.93 dBi and 8.73 dBi at 9.7 GHz and 12.7 GHz, respectively, which is not found in any previous designs. Thus, if used in arrays this design can find applications in very high gain requirements in the X band region.

Keywords Swastika · X band · Peak gain · Cross-polarization

1 Introduction

Owing to the massive miniaturization in the size of circuits and equipment, the importance of microstrip antennas is also increasing and so is the research in this field. The most significant part is the optimization of the design to suit a given application and this is based on several techniques outlined in the literature. Several methods of optimizing the parameters like bandwidth, VSWR (Voltage Standing

T. Gupta (✉)

Department of Electronics and Electrical Engineering, University of Petroleum and Energy Studies, 248007 Dehradun, India
e-mail: tgupta777@gmail.com

T. Kumar

Department of Electronics Instrumentation & Control Engineering, University of Petroleum and Energy Studies, Dehradun, India
e-mail: tkumar@ddn.upes.ac.in

R. Kamal

Department of Electronics and Communication Engineering, Dayananda Sagar University, Bengaluru 560078, Karnataka, India

Wave Ratio), radiation characteristics, gain, isolation in polarization exist [1–3]. The increase in substrate thickness and the lower dielectric constant lead to an increase in bandwidth, but the gain bandwidth product being a constant, the gain reduces. The increased level of surface waves when substrate height increases lead to a decrease in the gain of antenna. Introduction of slots in a simple rectangular patch in the Swastika configuration improves the radiation efficiency and thus the gain of the antenna. Various Swastika configurations have already been analyzed [4, 5] but have limited maximum gain to around 4dBi. The four slots in an orthogonal configuration typical of a Swastika lead to dual-frequency operation [6]. Cross-polarization leads to impurity in the desired polarization and thus needs to be reduced. Several techniques to lower the cross-polarization levels like using aperture coupling, proximity coupling exist [7]. Aperture coupling and proximity coupling are noncontact feed methods, which reduce contribution to broadside radiation from stripline feed which leads to increased cross-polarization levels. But this type of feeding method adds to the complexity of the design and also increases the height of the antenna which may not always be desired. Introduction of H-shaped slots, L-shaped slots, two separate feeds to feed the parallel slots, use of meta-material microwave absorbers to reduce coupling between the two orthogonal polarizations have also been practiced [8–10]. Modification in the ground plane like the introduction of slots and reduction of size can significantly change the radiation properties of an antenna [11]. Here, limiting the size of the ground plane towards the center to less than the dimensions of slots was observed to reduce cross-polarization levels.

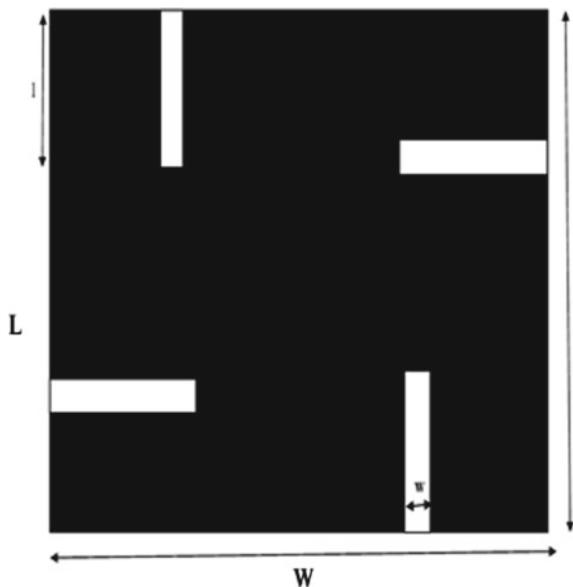
2 Proposed Antenna Design

Swastika is a typical square antenna with four slots organized in a mutually orthogonal pattern. In the design in Fig. 1, the metal patch is Swastika shaped with the side of square patch taken as 40 mm each. The four slots are each 1.5 mm wide and 10 mm long and are organized as shown in Fig. 1. The substrate lies between the metal patch and the ground plane. The height of the substrate is taken to be 1.588 mm and RT Duroid (Rogers) with dielectric constant ϵ_r of 2.2 is taken as the substrate material. The substrate is 25 mm wide and 25 mm long, with the substrate dimensions as 25 mm × 325 mm × 1.588 mm. The ground plane has been shortened to reduce cross-polarization and is 10 mm × 10 mm across the center of the antenna. This concentrates the electric field between the patch and the ground plane within this region, thus reducing coupling from the slots. The metal patch and the ground plane both use copper as the conducting material.

The antenna is fed by a coaxial feed going all the way from the ground plane to the metal patch and feeding it. Upon optimizing the feed location for best impedance match, the feed was connected at 3 mm from the center along the x-axis. Using microstrip feed adds radiation of its own (as it is made up of same material as patch radiator), therefore the simple coaxial feed with the ground plane as a reference has been put to use.

The antenna dimensions are given in tabular form.

Fig. 1 Proposed antenna design



Finite Element Method (FEM) based HFSS (High-Frequency Structure Simulator) software was used to design and optimize the proposed antenna with given dimensions and the results obtained were analyzed. Figure 2 shows the top view of the antenna as designed in HFSS, while Fig. 3 shows the bottom view with the reduced ground plane in square pattern. The top plane is the patch, ground plane at the bottom and the substrate between the patch and ground plane.

It is clearly seen in the top view that the metal patch has a Swastika configuration. The dark green portion in the bottom view is the ground plane, which has been reduced in size that it is away from the four slots. The light green circle is the coaxial feed used to excite the antenna. Upon simulation in HFSS, various results were analyzed and the antenna parameter like S_{11} , VSWR, gain, radiation efficiency, polarization levels, and radiation pattern were analyzed.

3 Results and Discussions

HFSS uses differential equation based FEM (Finite Element Method) to produce the results based on the design parameters used for the antenna. Initial designs with the full ground plane as the size of the substrate or even the metal patch showed significant cross-polarization levels. Thus, the ground plane size was reduced to suppress the mixing of polarization produced by the orthogonal slots. The feed position to get the suitable impedance matching was optimized using the parametric option available in Optimetrics in HFSS.

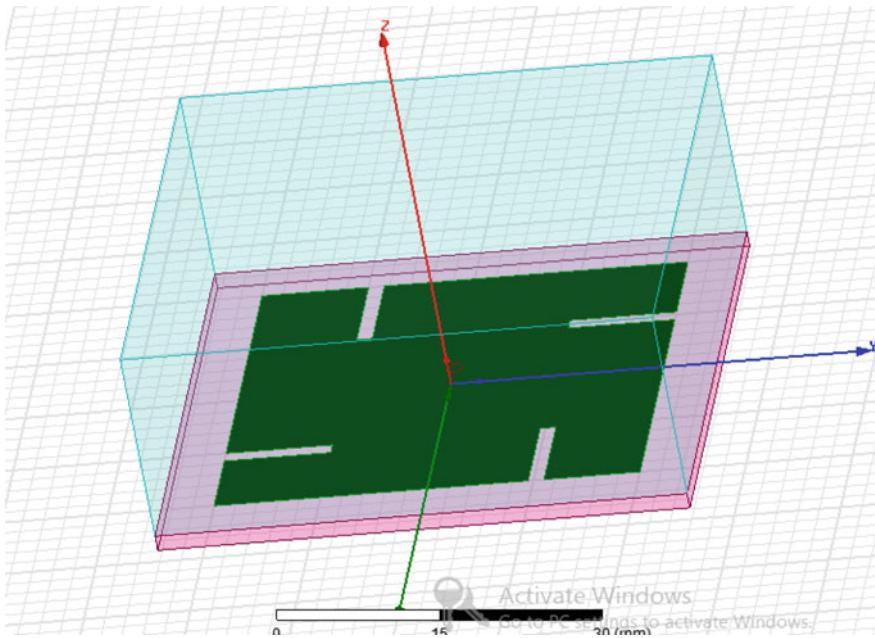


Fig. 2 Topview of the simulated design in HFSS

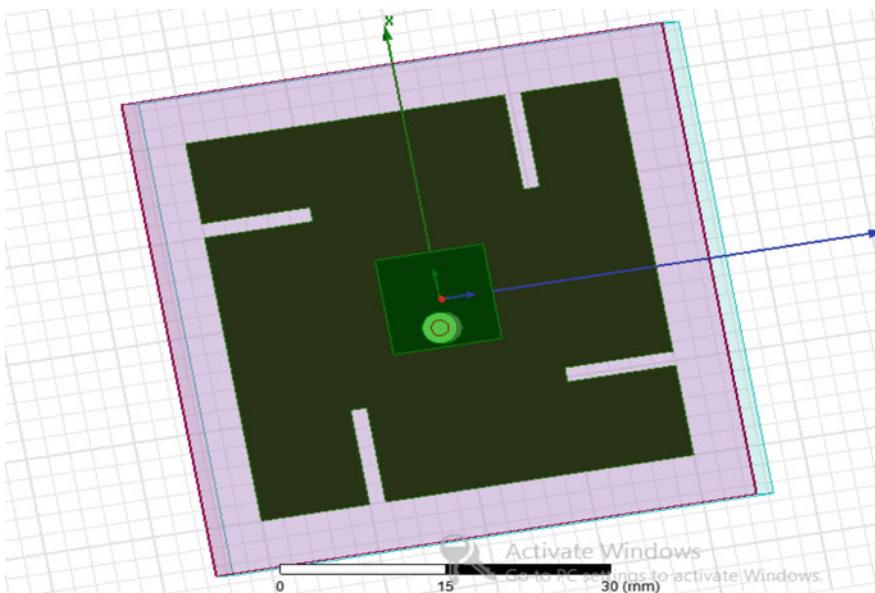
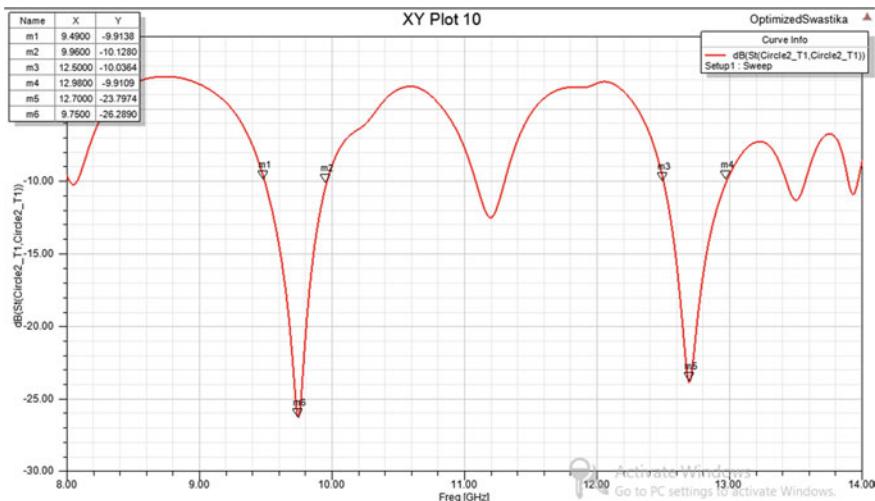


Fig. 3 Bottomview of the simulated design in HFSS

Table 1 Design parameters

S. no.	Design parameter	Value
1	Length of the patch (L)	40 mm
2	Width of the patch (W)	40 mm
3	Length of each slot (l)	10 mm
4	Width of each slot (w)	1.5 mm
5	Length of the ground plane	10 mm
6	Width of the ground plane	10 mm
7	Length of the substrate	50 mm
8	Width of the substrate	50 mm
9	Height of the substrate (h)	1.588 mm
10	Relative permittivity of the dielectric (ϵ_r)	2.2

**Fig. 4** S_{11} parameter versus frequency plot in HFSS

Thus, for the final design parameters as given in Table 1, we obtained two resonant frequencies of 9.7 and 12.7 GHz with 450 MHz -10 dB bandwidth from the S parameter plot as shown in Fig. 4. S parameter represents the reflection coefficient which is the ratio of the reflected to incident power at the given port. The value of S_{11} shows two notches of -26.2 and -23.7 dB at the resonant frequency of 9.7 GHz and 12.7 GHz, respectively. From the VSWR versus frequency plot in Fig. 5, the VSWR 2 bandwidth also corresponds to the -10 dB S parameter bandwidth at 450 MHz around each center frequency. VSWR shows a value of 1.1 at both the center frequencies and less than 2 over the bandwidth around these frequencies thus making it suitable for dual-band applications.

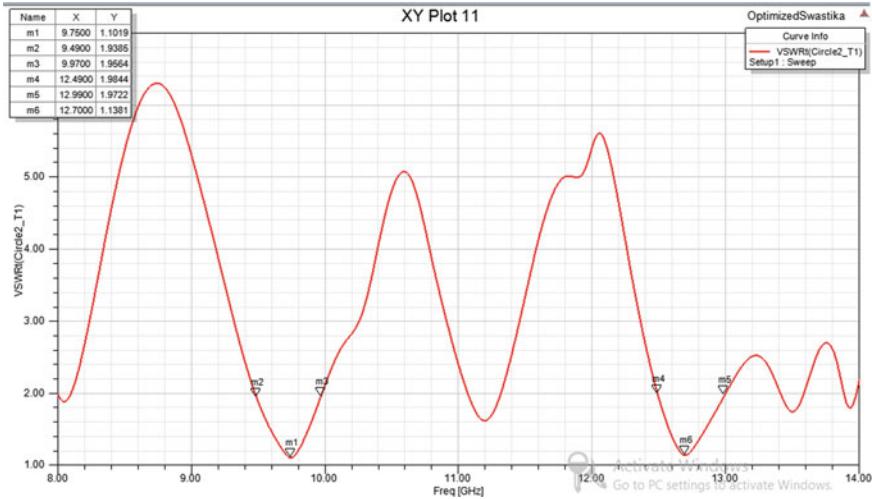


Fig. 5 VSWR versus frequency plot in HFSS

Table 2 Antenna parameters at 9.7 GHz

S. No.	Parameter	Value
1	Frequency	9.7 GHz
2	Max U	5.697 mW/sr
3	Peak directivity	5.9844
4	Peak gain	5.9349
5	Peak realized gain	5.9307
6	Radiated power	11.963 mW
7	Accepted power	12.063 mW
8	Incident power	12.071 mW
9	Radiation efficiency	0.99173
10	Front to back ratio	15.598
11	Decay factor	0.000000

From Tables 2 and 3, the values of other antenna parameters can be analyzed. For the resonant frequency of 9.7 GHz, the value of gain is obtained as 5.93 dBi and the radiation efficiency is 99.1%. Also, Table 3 shows the gain value at 12.7 GHz to be high at 8.73 dBi and radiation efficiency as 98.7%. Thus, the high value of gain at both the frequencies makes the antenna fall into the category of high gain antennas. When used in an array configuration, the gain can be further increased and thus make the antenna suitable for a greater number of applications.

The 3D radiation pattern for the two center frequencies of 9.7 and 12.7 GHz are shown in Figs. 6 and 7. The 2D radiation polar plot in Fig. 8 for 9.7 GHz and in Fig. 9 for 12.7 GHz shows the value of co-polarization and cross-polarization. In

Table 3 Antenna parameters at 12.7 GHz

S. no.	Parameter	Value
1	Frequency	12.7 GHz
2	Max U	8.21 mW/sr
3	Peak directivity	8.8475
4	Peak gain	8.7372
5	Peak realized gain	8.5557
6	Radiated power	11.674 mW
7	Accepted power	11.821 mW
8	Incident power	12.072 mW
9	Radiation efficiency	0.98754
10	Front to back ratio	48.579
11	Decay factor	0.000000

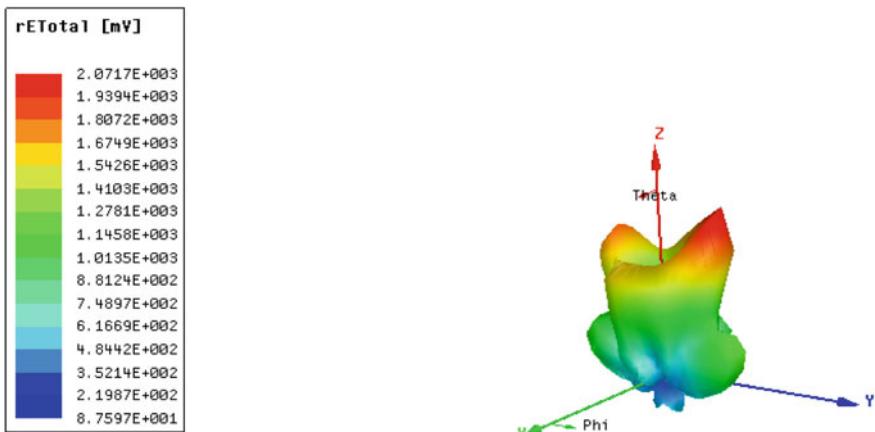


Fig. 6 3D radiation pattern at 9.7 GHz

Fig. 8, the inner circle represents the cross-polarization and it goes maximum 20 dB below the co-polarization levels with the least difference being -10 dB. At 12.7 GHz as shown in Fig. 9, the cross-polarization levels are even lower and better and they are at least 15 dB below the co-polarization levels.

4 Conclusion

The Swastika configuration antenna with a shortened ground plane designed in HFSS thus shows two center frequencies of 9.7 GHz and 12.7 GHz with a return loss of -21.9 dB and -22.7 dB, respectively, and peak gain of 5.93 dBi and 8.73 dBi,

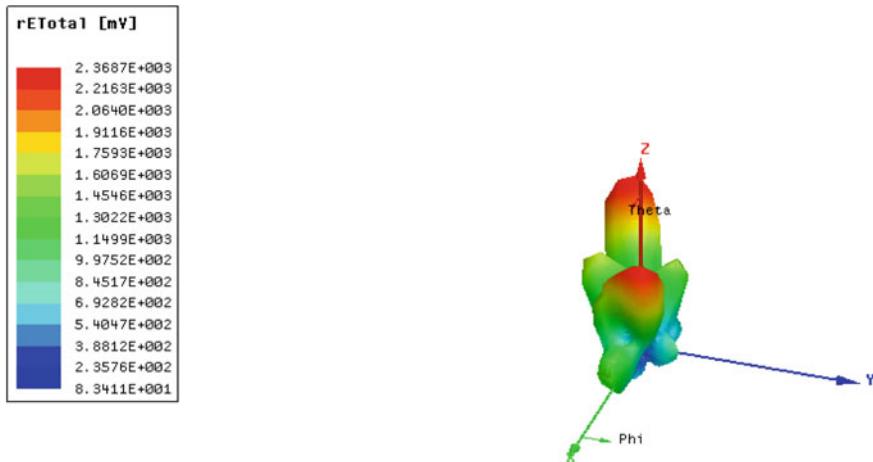


Fig. 7 3D radiation pattern at 12.7 GHz

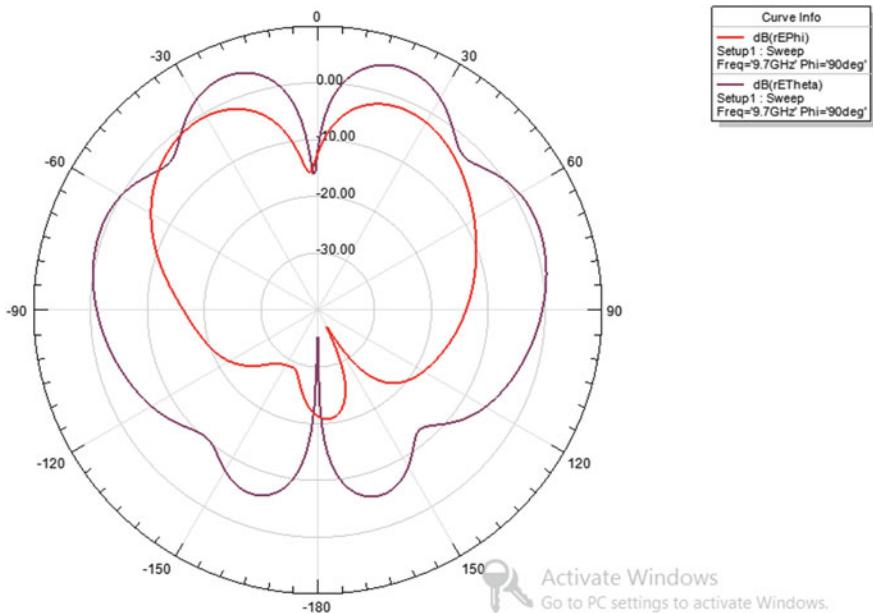


Fig. 8 2D radiation pattern at 9.7 GHz

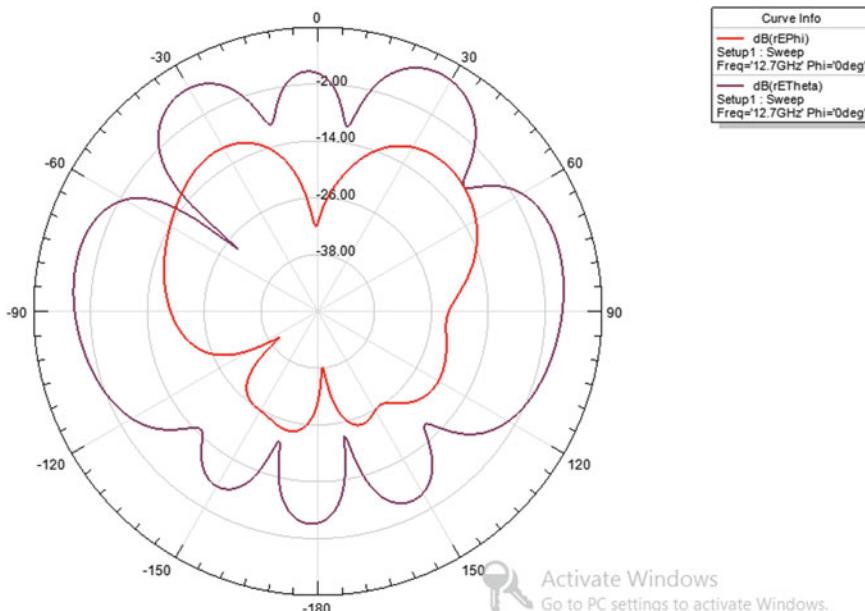


Fig. 9 2D radiation pattern at 12.7 GHz

respectively. Reducing the size of the ground plane was found to reduce the cross-polarization levels to at least 15 dB below the co-polarization levels. Such high gain antenna can find applications in RADARs operating in X Band when used in arrays to further increase the gain and effective area of the antenna. The dual-band characteristics make it suitable for HIPERLAN, automobile and defense applications also. Further optimization to increase the bandwidth and make it suitable for wideband applications can be undertaken.

References

1. Balanis, C.A.: *Antenna Theory*. Wiley, New York, pp. 727–734 (1982)
2. Bahl, I.J., Bharatia, P.: *Microstrip Antennas*. Artech House (1980)
3. Kumar, G., Ray, K.P.: *Broad Band Microstrip Antenna*. Artech House, pp. 1–21 (2003)
4. Samsuzzaman, M., Islam, T., Abd Rahman, N.H., Faruque, M.R.I., Mandeep, J.S.: Compact modified swastika shape patch antenna for WLAN/WiMAX applications. *Int. J. Antennas Propag.* (2014)
5. Rathor, V.S., Saini, J.P.: A design of swastika shaped wideband microstrip patch antenna for GSM/WLAN application. *J. Electromagnet. Anal. Appl.* **6**(03), 31 (2014)
6. Bhunia, S., Biswas, S., Sarkar, D., Sarkar, P.P.: Experimental investigation on dual-frequency broad band microstrip antenna with swastika Slot. *Indian J. Phys. Proc. Indian Associat. Cultivat. Sci. New Ser.* **81**(4), 495. Not Known (2007)

7. Gao, S., Sambell, A.: Dual-polarized broad-band microstrip antennas fed by proximity coupling. *IEEE Trans. Antennas Propag.* **53**(1), 526–530 (2005)
8. Li, M., Yang, H.L., Hou, X.W., Tian, Y., Hou, D.Y.: Perfect metamaterial absorber with dual bands. *Prog. Electromagnet. Res.* **108**, 37–49 (2010)
9. Wong, K.L., Tung, H.C., Chiou, T.W.: Broadband dual-polarized aperture-coupled patch antennas with modified H-shaped coupling slots. *IEEE Trans. Antennas Propag.* **50**(2), 188–191 (2002)
10. Chen, S.C., Wang, Y.S., Chung, S.J.: A decoupling technique for increasing the port isolation between two strongly coupled antennas. *IEEE Trans. Antennas Propag.* **56**(12), 3650–3658 (2008)
11. Chen, Z.N., See, T.S., Qing, X.: Small printed ultrawideband antenna with reduced ground plane effect. *IEEE Trans. Antennas Propag.* **55**(2), 383–388 (2007)

Review Analysis of Single-Electron Transistor



Raghav Bansal

Abstract In this paper, the results of the study of single-electron transistor have been summarized. The advancements in fabrication technology established by the well-known Moore's law had come to a standstill until the idea of using a single electron for operations came into effect. This single-electron transistor is a great revolution in research in nanotechnology. It uses a nanoscale island of charges called quantum dot for conductance and uses principles of 'Coulomb Blockade', single-electron tunneling effect. This allows miniaturization of circuits, reduced power consumption, high operating speed. These characteristics permit its application in various areas such as memory device industry, metrology. However, it has some fabrication issues which have not been fixed yet. At room temperature, a sub-nanometer island size is required for proper functioning and obtaining a good control over nanoparticle size is challenging. A comparison between SET and other conventional transistors has also been made.

Keywords Coulomb blockade · Single-electron transistor · Single-electron tunneling effect · Quantum dot

1 Introduction

The transistor is a basic building block of all modern electronic devices. It was invented in 1947 at Bell Labs and has emerged to be one of the most important inventions of the twentieth century. The technology used in transistors has been revolutionized ranging from BJTs to JFETs to MOSFETs using different semiconductors including silicon, germanium. A few decades after the invention, Moore's law came to be known well which established that the number of transistors in a dense integrated circuit double about every 2 years. This means that for the same chip area the complexity of circuits is increased as more and more transistors are added to the

R. Bansal (✉)

Department of Electrical and Electronics, BITS Pilani, Pilani Campus, Pilani, Rajasthan, India
e-mail: raghav17bansal@gmail.com

chip. This comes with added benefits such as high speed and low power consumption. This can be interpreted as the size of the transistor to be reduced to half every 2 years. Since then, the scientists have been involved in miniaturizing the circuits, till the idea about having control over a single electron gave rise to single-electron transistor. This device can bring about revolutionary changes if it is implemented in the real world and researchers are already doing continuous research to bring about change in current era.

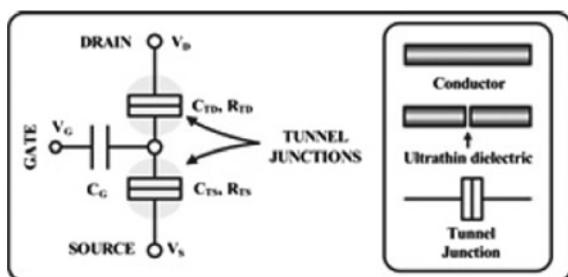
2 SET Principle and Working

On a macroscopic scale, the transistor consists of millions of electrons and it is simply not possible to control a single electron. However, on a nanoscale, the discreteness of current shows up and it becomes possible to gain access over an electron and control its movement. Here, these lengths are comparable to the system dimensions.

A single-electron transistor is a three-terminal device with the drain, source, and gate serving as its terminals made of metal [1–3]. In between the source and drain, there exists an island of electrons called the quantum dot which is involved in the conduction as opposed to a conducting channel in conventional transistors. A quantum dot is a mesoscopic system and with the presence or absence of an electron the electrostatic energy changes. This dependence on a single particle can be controlled using the gate voltage, which allows the current to be treated as discrete and not a continuous flow of electrons.

In between the source and dot and between the dot and the drain, there exists a tunnel junction (Refer Fig. 1) which acts as an insulating barrier which is described by capacitance and resistance. The electrons tunnel through the barrier one by one and this is treated as a discrete current. Since current is associated with resistance, the junction has a resistance which indicates a loss. In classical thermodynamics, an electron cannot pass through the insulating barrier. However, in quantum mechanics, there is a small probability for the electron to cross the barrier. Here, the island of electrons can be treated as a capacitance as it stores charge and therefore energy. The capacitance of the dot is expressed as the capacitance of a spherical capacitor as

Fig. 1 SET schematic [4]



$$C = 4\pi \varepsilon R \quad (1)$$

where R is the radius of the capacitor. The increase in electrostatic energy required to bring an electron from source to the dot is

$$E_C = \frac{e^2}{2C}. \quad (2)$$

where C is the capacitance of the quantum dot. This energy serves as the potential barrier for electron transfer known as the Coulomb Blockade energy. This is the repelling energy that an electron feels due to the previous electron present. It is inversely proportional to the capacitance, which in turn varies directly according to its radius, and since the radius is very small for a nanoscale device so the capacitance is very small and the energy becomes very large. This is why the electrons cannot pass through the junction simultaneously and hence this energy acts as a suppression for electron movement in the device, and hence it is called the Coulomb Blockade energy. There are ways to overcome this energy:

- (a) If at a particular temperature T, the thermal energy $E_T = kT$ becomes greater than the Coulomb blockade energy.
- (b) If the externally applied voltage V becomes greater than V_{th}

$$V_{th} = \frac{e}{2C}. \quad (2)$$

where V_{th} is known as the threshold voltage, which is defined as the critical voltage which if applied externally will build up sufficient energy to just counter the Coulomb blockade energy so that an electron can easily pass through the tunnel junction (insulating barrier) which can constitute current. This means that no current will flow across the junction if the applied voltage is less than the threshold voltage. But if the voltage is greater than V_{th} , the energy will exceed the Coulomb blockade energy and a current will flow.

Hence the current will not flow in these conditions:

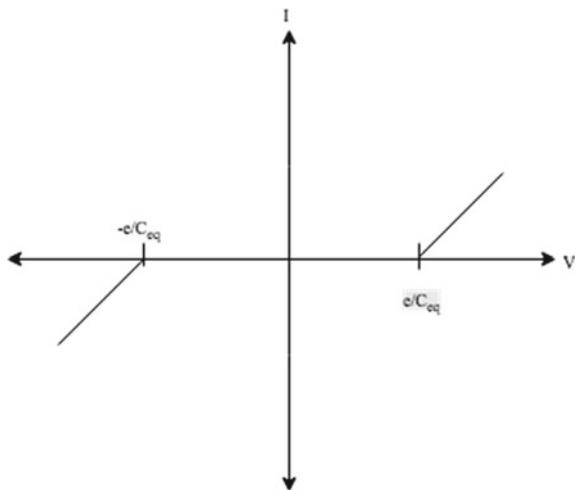
- (1) The external applied voltage must be less than e/C .
- (2) The heat energy should be less than the Coulomb blockade energy, i.e., $k_B T < \frac{e^2}{2C}$ otherwise, the electron shall pass through the quantum dot.
- (3) According to Heisenberg's Uncertainty Principle [4]

$$\Delta E \Delta t > \hbar \quad (3)$$

where $\Delta t = R_T C$ and $\Delta E = \frac{e^2}{C}$. Thus

$$R_T > \hbar/e^2 = 25813\Omega [6]. \quad (4)$$

Fig. 2 I-V characteristics for symmetric junctions [5]



3 I-V Characteristics

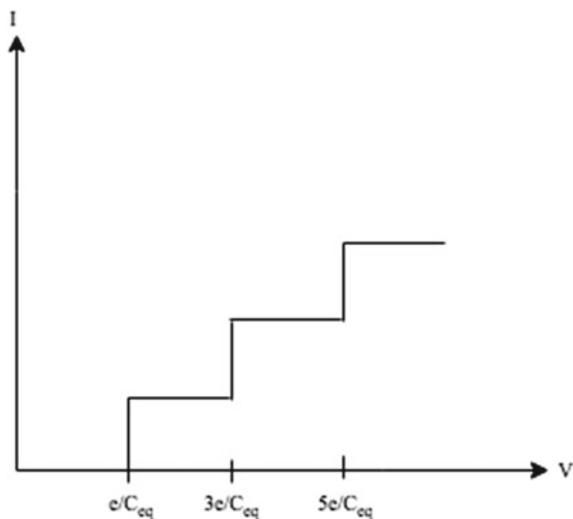
There is no current in a single-electron transistor if the applied voltage is less than the threshold voltage so that the energy becomes less than the Coulomb blockade energy. But if the externally applied voltage is greater than the threshold voltage, the current flow is recorded as the Coulomb blockade effect is removed (Refer Fig. 2).

For an asymmetric junction, electron enters through one junction and leaves through the other due to high resistance. Another electron leaves and exits rapidly. If the external bias is increased, the number of electrons on the dot is increased. The current appears to increase in the pattern of a staircase and the I-V characteristic is called ‘Coulomb Staircase’ (Refer Fig. 3).

4 Advantages of SET

- Low power consumption,
- High speed,
- Increased Sensitivity,
- Reduced size,
- Easy co-integration with the conventional CMOS circuits,
- Simple principle of operation, and
- Feature of reproducibility.

Fig. 3 I-V characteristics of an asymmetric junction [5]



5 Comparison Between SET and Conventional Transistors

5.1 SET and MOSFET

The physical structure of single-electron transistors is almost similar to MOSFETs. In spite of this, SETs and MOSFETs have a different principle and working such as SETs have an insulating barrier called the tunnel junction instead of p-n junction of the MOSFETs. The former also has an island of charges called quantum dot instead of conducting channel in MOSFETs.

In single-electron transistor when the electron overcomes the effect of Coulomb blockade it gets transferred from the source region to the drain region singly by single-electron tunneling whereas in MOSFET drain current is constituted by a large number of electrons crossing the channel simultaneously. As an electron reaches the quantum dot, it experiences electrostatic repulsion due to the presence of other negative charges. Thus, the number of electrons can be controlled and the drain current can be varied. The I-V characteristic is recurring denoting a finite current at particular voltages. Unlike FETs where the magnitude of current depends on the number of electrons, in SETs, it is independent of the number of electrons and on their Fermi velocity.

The common principle in both the devices is the electrostatic effect that governs their working. In SETs, it is due to the Coulomb blockade effect and due to large resistances between source and drain separated from the channel, the electron cannot cross the tunnel junction easily unlike ideal FETs.

5.2 SET and CMOS

The SET-based logic circuits working upon the tunnel junction to implement the SET as a switch equivalent of the MOS transistor. This is advantageous as existing CMOS transistor designs can be easily used to apply SET technologies. The drawbacks include the current transfer through an open transistor depends on a large number of individual electrons passing through the tunnel island junction. This process is slower than the transport of single electron and therefore the approach that uses the CMOS style should not use the SET technology at its fullest.

6 Applications of SET

(1) Memory Device

A single-electron memory is a very interesting application of the SET. There are many processes being developed for the same and it is emerging as a very competitive area. However, the limitations on the fabrication of such a small size device don't allow for this.

(2) Single-Electron Spectroscopy

This provides with the possibility to study the dynamics in atoms and the energy level distribution in nanoscale objects. There are two ways to accomplish this. One is to take the quantum dot as an island of electrons that is capacitively coupled to the transistor and the other is to have a low applied voltage and take the quantum dot directly as an island of charges and measure gate voltages with increasing conductance.

(3) Ultrasensitive Microwave Detector

The quantum dot is connected to the terminals of the transistor through tunnel junctions and capacitive gate. When low bias voltages are applied, the device can work only through photon-assisted tunneling due to Coulomb blockade. After this is done, the particle remains trapped in the island as it takes comparatively longer time to pass through. Since it makes the device current on, two electrons tunnel the junction at a time instead of one and thus, the device becomes highly sensitive.

(4) Temperature Standards

$\Delta V = \frac{5.44Nk_B T}{e}$ has been found to be astonishingly stable with reference to any dissimilarity in the array parameters except the considerable spread in the tunneling resistances, on a theoretical note. As the basic constants have been calculated with high accuracy, the arrays can be easily used for measuring absolute temperatures. A single array may give a high accuracy; however, it may be limited to less than one decade in variations of temperature. This range may vary with arrays having a distinct size of the dots. This gives a promising way of finding the absolute standard of temperature over a large range from a combination of different circuits developed on one chip. This method will certainly

be a challenge for conventional temperature standards, however, it is still in its infancy and it will require time to develop to its full capacity.

(5) **Supersensitive Electrometer**

The development of insulating barriers in single-electron transistors is not yet easy at the moment as more development is needed in its technology. Still, they have proved to be successful supersensitive electrometers in certain physical experiments due to their high sensitivity like they have been used to make undeniable observations of the parity effects in semiconductors. They comprise many electron trapping centers and systems efficient in producing random low-frequency modifications in conductance of the insulating barrier. Changes in the transistor have enabled to prove the presence of fractional charge excitations in fractional quantum Hall effect. It has also been used to measure very low DC currents and single-electron effects in lone charge devices.

(6) **Detection of Infrared Radiation**

Another important application of SETs is in the detection of infrared radiations at room temperature. The sensitivity of the device can be managed and the wavelengths can be detected over a wider range through the use of excited charges. It works whenever it comes in the influence of an infrared signal which triggers the conduction band electrons. A silicon insulator channel raises the charge as the number of excited electrons increases. Further, a polysilicon gate switches off the transistor and creates a node where electrons can enter by forming an energy barrier. Thus, a current is recorded whenever an electron with energy higher than that of the energy barrier enters into the storage node.

(7) **Charge State Logic**

A charge state logic device, where the existence/nonexistence of an electron represents one bit of data at definite quantum dots, is used to resolve the leakage current issue in SETs. Since there is no DC current at static state, there is no current and power.

7 Limitations on SET Implementations

(1) **Room Temperature**

This is a major issue that is not allowing SETs to be implemented at room temperature as the thermal energy $k_B T$ adds on to the electron energy and may disrupt the energy threshold required for electron flow. Small variations cause large variations in the island shape and lead to unacceptable variations in spectrum.

(2) **Background Charge**

The background charge produces random results. A unipolar impurity in the tunnel junction separates the environment into dipoles and induces an image charge on the quantum dot. This disrupts the electrostatic energy of the dot and other conditions such as threshold voltage changes.

(3) Co-tunneling

The speed of the process decreases due to tunneling of more than one electron at the same time. It is viable that more than one electron tunnel through different insulating barriers at the same moment as one coherent process.

(4) Lithography Techniques

A requirement for the proper working of single-electron transistor is that $E_C \sim 100k_B T$. This means a nanoscale island for room temperature operation. But practically using present VLSI techniques, this is very difficult. Moreover, even if this is fabricated, the shape of the island shall not be perfectly regular.

(5) Outside Environment Linking with SETs

There are two ways—one is to combine SET technology with the existing MOS-FET devices which will make the integration density attractive. The other is to not use the wire, but linking the clusters using static electronic forces between them to form circuits linked by cluster called quantum cellular automata (QCA). No wire is required for such integration and results in high-density memory.

8 Conclusion

SET have already proved their value as a tool in scientific research. They have an island of charges (quantum dot) instead of the channel which provides for conductance and work on the phenomenon of ‘Coulomb Blockade’. The insulating barriers between the source and the dot, and between the dot and the drain called the tunnel junction can be represented as a resistance (depends on tunneling of the electron) together with capacitance (depends on the size of the nanoparticle). The resistance is determined by tunneling of electron and the capacitance depends on the size of the nanoparticle. If the applied voltage is greater than V_{th} , the energy becomes greater than the suppression energy E_c due to which the electron is able to cross the potential barrier and current flows.

SET provides many advantages over conventional devices like increased speed, low power consumption, high sensitivity and many more. These benefits enable the SETs to be applied in various areas like metrology, spectroscopy, in supersensitive electrometers, as a memory device, detection of radiations.

However, in the present scenario, SET comes with many limitations. It is very difficult to operate the device at room temperature as it causes disruptions in its operations. Also, the device is difficult to fabricate using current techniques and there is a problem of power leakage causing excessive heat from transistors in close proximity.

Not to mention, the researchers are finding ways to overcome these difficulties. The single-electron transistors will surely be able to replace the conventional transistors in the near future.

References

1. Kumar, O., Kaur, M: Single electron transistor: applications and problems. *VLSICS* **1**(4), 24–29 (2010)
2. Kumar, A., Dubey, D.: Single electron transistor: applications and limitations. *AEEE* **3**(1), 57–62 (2013)
3. Aguiam, D., Vincé, O.: A Brief Introduction to Single Electron Transistors (2011)
4. Sarmiento-Reyes, A., Castro González, F.J., Hernández-Martínez, L.: A methodology for simulation of hybrid single-electron/MOS transistor circuits. *26*(2) (2013)
5. Gupta, M.: A study of single electron transistor. *IJSR* **5**(1), 474–479 (2016)
6. Singh, G.P., Raj, B.: Single Electron Transistor Theory: A Review. *ICE-CCN*, pp. 95–99 (2013)

Author Index

A

- Aayushi, 773
Agarwal, Shivi, 393, 895
Aggarwal, Alok, 707
Agrawal, Sunil, 85
Agrawal, Anjali, 551
Ahmed, M. A., 935
Ahuja, Neelu Jyoti, 425, 707
Ambigavathi, M., 755
Anand, Karan, 313
Anishchenko, Oleksandr, 325
Arora, Monika, 483
Ashish, Shubham, 463
Awasthi, Neha, 233
Ayare, Natasha Sanjay, 773
Ayush, 599

B

- Babel, Mudit, 661
Baghel, Abhishek, 671
Balalayeva, Elena, 325
Bansal, Hari Om, 523, 807
Bansal, Kritika, 647
Bansal, Raghav, 975
Basha, Mudasar, 851
Bhadka, Harshad, 469
Bhardwaj, Sanjay, 377
Bhardwaj, Shantu, 345
Bharti, Vishal, 637
Bhaskar, Basant Singh, 765
Bhaskar, Satyam, 807
Bisht, Amandeep, 385
Bisht, Mangal Singh, 95

C

- Chacko, Saji, 163, 213
Chaitra, P. G., 781
Chakrabarti, Prasun, 37, 283
Chamoli, Vivek, 105
Chandini, S. B., 773
Chaudhari, Kinjal, 819
Chauhan, Piyush, 627
Chopra, Akanksha Bansal, 483
Choudhary, Yogendra, 833
Choudhury, Sushabhan, 883
Chowdary, Vinay, 883
Chukka, Ramesh Babu, 875

D

- Dahiya, Ratna, 333
Deb, Arindam, 265
Debbarma, Praseed, 265
Deepthi, V., 781
Devalla, Vindhya, 765
Devesh, Mishra, 367
Dharavath, Abhiraj, 599
Dhindsa, Anaahat, 175
Dixit, Anurag, 857
Dixit, Veer Sain, 353, 483
Dobhal, Daksh, 589
Dogra, Ayush, 85
Dutta, Sagnik, 727
Dutta, Subhankar, 661
Dwivedi, R. P., 1

G

- Gandigudi, Niranjana, 313

- Gandotra, Neeraj, 1
 Garg, Sagar, 463
 Ghosh, Pratik, 405
 Goel, Arnav, 589
 Goel, Naveen, 213
 Goel, Somya, 671
 Goutham, Veerapu, 609
 Gowri, R., 833, 921
 Goyal, Bhawna, 85
 Goyal, Piyush, 627
 Gupta, Mukul Kumar, 531, 627, 913
 Gupta, Rajeev, 1
 Gupta, Shweta, 415
 Gupta, Tanisha, 965
- H**
 Haridas, S. L., 697
 Harigovindan, V. P., 609
 Hariprasad, S. A., 415
 Harsoda, Harshad, 445
 Hasan, Naimul, 405
 Hinduja, Akshay, 243
 Hurkovska, Svitlana, 325
- I**
 Indumathi, P., 735
- J**
 Jagabar Sathik, M., 661
 Jain, Animesh, 807
 Jain, Anuj, 37
 Jain, Arpit, 841
 Jain, Mohit, 747
 Janani, R., 275
 Jately, Vibhu, 559
 Jayamangala, S., 619
 Joshi, Upasna, 719
 Jyothi, N., 799
- K**
 Kamal, Rajeev, 965
 Kant, Agrawal Krishna, 367
 Kartikeya, Kumar, 589
 Kashyap, Kamal, 435, 903
 Kathiravan, K., 13
 Kaundal, Vivek, 727, 765, 913
 Kaur, Amanpreet, 53, 113
 Kaur, Arashpreet, 113
 Kaur, Chamandeep, 385
 Kaur, Gagandeep, 53
 Kaur, Manpreet, 589
 Kaur, Parminder, 203
 Kaushik, Shiva, 157
 Kaushik, Keshav, 791
- KaviPriya, M., 735
 Kene, Pragati, 697
 Keshav, Kishore, 453
 Khattri, Vipin, 73, 123
 Khetwal, Nishi, 765
 Kim, Dong-Seong, 377
 Kothi, Nakshatra, 37
 Kothiyal, Alok Darshan, 95
 Kotibhaskar, Nakul, 647
 Kukhar, Volodymyr, 325
 Kumar, Adarsh, 707
 Kumar, Adesh, 415, 765, 851, 913, 945
 Kumar, Amit, 425, 637
 Kumar, Anil, 95, 435, 903
 Kumar, Ch. Sumanth, 875
 Kumar, D., 619
 Kumari, Ch. Aruna, 225
 Kumari, N. Prasantha, 225
 Kumar, Megha, 719
 Kumar, Nishant, 841
 Kumar, Ravinder, 523
 Kumar, Ravindra, 719
 Kumar, Roushan, 913
 Kumar, Sanjeev, 47
 Kumar, Shashwat, 727
 Kumar, Shubham, 841
 Kumar, Shukla Narendra, 367
 Kumar, Sunil, 95
 Kumar, Tarun, 833, 965
 Kumar, Vineet, 295
 Kuppast, V. V., 689
 Kurbet, S. N., 689
 Kurien, Caneon, 313
- L**
 Laxkar, Pradeep, 37, 283
- M**
 Madhavi, D., 799
 Madhu, Charu, 175
 Madhuri, G. M. G., 225, 957
 Mahobia, Ravi Prakash, 213
 Malik, Monu, 333
 Manda, Rajarao, 921
 Manoj, Manuja, 453
 Markov, Oleg, 325
 Marwaha, Anu, 133
 Mathur, Trilok, 393, 895
 Mathur, Vishal, 321
 Maurya, Shubham, 747
 Meera, C. S., 531, 627
 Mishra, Jagriti, 531
 Mitra, Urvi, 569
 Mittal, Udayveer, 765

Mohammad, Salauddin, 619, 851
Mondal, Amit Kumar, 727
Monica, Ch., 225
Mukherji, Manishita, 321
Mukherji, Rana, 321
Mukhija, Pankaj, 647
Mungra, Dhara, 551

N

NagaRaju, Vonteddu Shanmukha, 25
Nag, Atanu, 405
Nagrath, Preeti, 671
Nandanwar, Kunal, 807
Nayak, Sandeep Kumar, 73, 123
Negi, Aradhana, 203
Nida Susan, A. K., 773
Nithyanandam, S., 513

P

Pachauri, Nikhil, 747
Pachpor, Siddharth, 661
Pal, Tamistha, 265
Panda, S. N., 345
Pandey, Alok Kumar, 463
Pandey, Jitender, 105
Pandey, Manju, 243
Pandey, Rahul, 163
Pandey, Surya, 255
Pandian, R., 735
Pant, Vishal, 833
Panwar, Shailesh Singh, 497
Patnaik, M. R., 799
Prajapati, Payal, 819
Prakash, Rishi, 105
Prasanthi kumari, N., 957
Praveen Kitti, B., 957
Prema, N. S., 773
Pritam, Kocherlakota Satya, 393
Priya, Rashmi, 405
Prysiazhnyi, Andrii, 325
Purkait, Sonal, 163

R

Radha, R., 13
Raiwani, Y. P., 497
Raja Kumar, R., 735
Rajini, S., 781
Ramalla, Isaac, 935
Ramli, Muhammad Rusyadi, 377
Rao, Rayapudi Srinivasa, 25
Rathore, Gunjan, 175
Ravindra, Manam, 25

Reddy, S. R. N., 195
Rehman, B. Khaleelu, 619, 851
Rekha, G., 539
Rijuvana Begum, A., 513
Roy, Anwesha, 265
Roy, Kousik, 405

S

Saeed, Fatek, 857
Sahirov, Yurii, 325
Sahu, Lalit kumar, 163
Saini, Devender Kumar, 1
Saini, Raj Kumar, 1, 63, 903
Sappal, Amandeep Singh, 143
Saraswat, Sagar, 463
Sarkar, Prasanta, 303
Satheeswaran, C., 735
Seetha, R., 619
Shah, Dikshan N., 469
Shah, Kunal, 283
Shah, Shashvat, 445
Shamneesh, Sharma, 453
Shareef, Zeenat, 195
Sharma, Abhinav, 833
Sharma, Anupam, 865
Sharma, Apoorav Maulik, 85
Sharma, Ashwani, 1, 63
Sharma, Himanshu, 913
Sharma, Madhu, 865
Sharma, Ram Mohan, 935
Sharma, Ritu, 213
Sharma, Shalini, 63
Sharma, Sukesha, 233
Sheth, Jeel, 445
Shimi, S. L., 945
Shukla, Aasheesh, 157, 185
Singhal, Ravi, 707
Singh, Deepak Kumar, 73, 123
Singh, Dheerendra, 523
Singh, Harmeet, 143
Singh, Inder, 63
Singh, Juhi, 185
Singh, Kamalpreet, 791
Singh, Krishna Veer, 523
Singh, Lairenlapam Joyprakash, 303
Singh, Madhusudan, 589, 599
Singh, Preeti, 47, 385
Singh, Shikha, 579
Singla, Anshu, 133
Sisay, Asefa, 559
Sreenath, N., 539
Sridharan, D., 755

Srivastava, Ajay Kumar, 313, 727
Srivastava, Aprajita, 671
Srivastava, Garima, 569
Srivastava, Tanya, 579
Suchit, Yadav Ram, 367
Suresh Kumar, P., 957
Suri, Amar Raj Singh, 903
Swarnakar, Jaydeep, 303

T

Talikoti, Basavaraj, 689
Taruja, Pande, 367
Thakkar, Ankit, 445, 551
Thakur, Robin, 1, 63, 435, 903
Thakur, Sourav, 765
Thapa, Sashank, 435
Thirunavukkarasu, I., 275
Thrlochan Sharma, P., 883
Towhidul Islam Nayim, S. M., 435

Tyagi, Aayushi, 671
Tyagi, Amit Kumar, 13, 539

U

Udupa, Pradeep, 513
Upadhyay, Shivansh, 727

V

Valluru, Sudarshan K., 589, 599
Veeramalla, Vineeth, 531
Venkata Siva Reddy, K., 851
Verma, Anjani Kumar, 353
Verma, Piush, 1
Verma, Varnita, 531, 627
Vidyashree, K. P., 781
Vig, Renu, 85
Vishnoi, Priyanshi, 945
Vora, Kevin, 445