Lecture Notes in Networks and Systems 11

Swapan Bhattacharyya
Sabyasachi Sen
Meghamala Dutta
Papun Biswas
Himadri Chattopadhyay *Editors*

Industry Interactive Innovations in Science, Engineering and Technology

Proceedings of the International Conference, I3SET 2016

Springer

Lecture Notes in Networks and Systems

Volume 11

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

The series "Lecture Notes in Networks and Systems" publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Advisory Board

More information about this series at http://www.springer.com/series/15179

Swapan Bhattacharyya · Sabyasachi Sen
Meghamala Dutta · Papun Biswas
Himadri Chattopadhyay
Editors

Industry Interactive Innovations in Science, Engineering and Technology

Proceedings of the International Conference, I3SET 2016

Springer

Editors
Swapan Bhattacharyya
Department of Electronics and Communication
JIS College of Engineering (JISCE)
Kalyani, West Bengal
India

Sabyasachi Sen
Department of Physics and Nanoscience and Technology
JIS College of Engineering (JISCE)
Kalyani, West Bengal
India

Meghamala Dutta
Department of Biomedical Engineering
JIS College of Engineering (JISCE)
Kalyani, West Bengal
India

Papun Biswas
Department of Electrical Engineering
JIS College of Engineering (JISCE)
Kalyani, West Bengal
India

Himadri Chattopadhyay
Department of Mechanical Engineering
Jadavpur University
Kolkata, West Bengal
India

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-10-3952-2 ISBN 978-981-10-3953-9 (eBook)
DOI 10.1007/978-981-10-3953-9

Library of Congress Control Number: 2017934629

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Foreword

International Conference on Industry Interactive Innovations in Science, Engineering and Technology (I3SET) 2016 is scheduled to be held during October 25–26, 2016. The proceedings of the conference will be published in *Lecture Notes in Networks and Systems* (LNNS), Springer. This is the fourth consecutive international conference in successive years that the Institution has been hosting under Technical Education Quality Improvement Program (TEQIP). Previously, our Institution hosted International Conference in Computation and Communication Advancement (IC3A) in the year 2013; the proceedings of which were published in *McGraw Hill*, International Conference on Non-Conventional Energy (ICONCE) 2014 published in *IEEE Xplore* and International Conference on Advancement in Medical Electronics (ICAME) 2015 published in *Lecture Notes in Bioengineering, Springer*.

I3SET is convergently focused on encouraging and escalating innovations in a demand driven way as desired by industry. From very inception of hosting this conference, the idea was to make a tripartite collaboration between academia, industry and research in a think-pair-share mode.

Influenced and inspired by the global trends in expansive application and intensive research opportunities in industry interactive innovation, the institution

has been expediting and exploring avenues in science, engineering, technology and management, with dual objective of strengthening research & development activities as well as industry institute partnership. I3SET 2016 is a collective initiative focused on specialized as well as interdisciplinary research opportunities.

I3SET provides versatile domains, viz. nonconventional energy & advanced power systems, nanotechnology and applications, pattern recognition and machine intelligence, digital signal and image processing, modern instrumentation, control, robotics and automation, civil engineering and structural design, real-time and embedded systems, communication and devices, advanced optimization techniques, biotechnology, biomedical instrumentation and bioinformatics extended to outcome-based education. In spite of the versatility, the objective of the Conference is convergent on annexing innovations and applications to visualization and conceptualization of emerging topics uplifting "ideas embedded with innovations".

Dr. Somsubhra Gupta
Convener, I3SET 2016
Associate Professor and Dean—Academic Affairs
JIS College of Engineering (An Autonomous Institution)
Kalyani, West Bengal,
India

Preface

I3SET 2016 focused on to bring together innovative academicians and industrial experts in the field of science, engineering and technology to a common forum. The conference is sponsored by *Technical Education Quality Improvement Programme (TEQIP)*.

The idea of the conference is for the scientists, scholars, engineers and students from the universities all around the world and the industry to present ongoing research activities on the relevant topics and hence to foster research activities between the universities and the industries. I3SET provides a valuable opportunity for them to exchange their ideas face to face together. I3SET is very special for its strong organization team, dependable reputation and wide sponsors from all around the world. It will bring you an unexpected harvest. In the future, you may be a member of our big family. Therefore, it will be a good opportunity to share your innovative thoughts with other renowned researchers in respective research fields across the world.

In this International Conference, a total of 79 papers were accepted for presentation against submission of 139 papers overall. All presented papers will be available online in Springer Link. We have the privilege to invite research-based papers, articles on the broad theme of the conference in relevant track areas from esteemed Institutions in the country and abroad. English is the official language of the conference. The conference proceedings are published in Springer—Lecture Notes in Networks and System (LNNS).

The I3SET 2016 has the following announced Tracks:

Track 1: Nonconventional Energy and Advanced Power Systems

Subtrack 1.1 Solar Energy:
Solar photovoltaic technology, Solar-thermal energy, Solar energy technology, Solar architecture
Subtrack 1.2 Hydro Energy:
Sea Energy, Wave energy, Wave energy system, Tidal energy, Tidal sea current, Turbo machinery, Mini hydro
Subtrack 1.3 Bio Energy:
Bio Fuels, Energy from waste, Bio mass & Bio gas
Subtrack 1.4 Hydrogen Energy:
Hydrogen and energy storage and transportation, Hydrogen and fuel cell, Carbon sequestration
Subtrack 1.5 Energy from Depth of Earth:
Geothermal energy, Ocean-thermal energy
Subtrack 1.6 Wind Energy:
Wind turbine, Wind mill, Wind firm, Wind potential assessment (including offshore)
Subtrack 1.7 Energy & Solar Passive Architecture:
Energy conservation in built in system, Solar heating & cooling, Zero energy building
Subtrack 1.8 Energy Conservation Measures:
Optimization techniques, Application of artificial intelligence in energy system, Grid interactive system, Micro grid and smart grid, Energy storage, Fuel cell, Electric and hybrid vehicles, Efficiency in irrigation

Track 2: Nanotechnology and Applications

Subtrack 2.1 Nanomaterial Particles and Applications:
Nanofabrication technologies, Carbon nano-tubes and grapheme technologies, Nanocomposites, Characterization and properties of nanomaterials, Simulation Modelling of nano-materials
Subtrack 2.2 Nano Science and Technology:
Nano electronics, Nano medicine, Medical nanotechnology
Subtrack 2.3 Molecular Nanotechnology:
Nano robotics, Advanced grapheme science, Metal nanocrystal, Nano molecular material design
Subtrack 2.4 Nanotechnology for Energy Systems:
Nano-optimization for fuel cell and solar cell, Nano-membrane and nanosieve

Subtrack 2.5 Emerging Areas of Nanotechnologies:
Sprintonics, Nanomagnetism, Nanoprinting, Nanopacking

Track 3: Pattern Recognition and Machine Intelligence

Subtrack 3.1 Biometrics:
Biometric systems and applications, Multi-biometrics, Forensic biometrics and its applications, Fingerprint recognition, Face recognition, Iris recognition, Soft biometrics
Subtrack 3.2 Human Computer Interaction:
Human computer interaction, Display hardware, Character and text recognition, Handwriting recognition, Graphics recognition, Human body motion and gesture based interaction, Affective computing, Facial expression recognition, Group interaction: analysis of verbal and non-verbal communication, Gate recognition, Speaker recognition
Subtrack 3.3 Computational Intelligence:
Fuzzy computing, Rough computing, Granular computing, Evolutionary computing, Neural computing, Case based reasoning
Subtrack 3.4 Machine Learning:
Statistical, Syntactic and structural pattern recognition, Machine learning and data mining, Symbolic learning, Active and ensemble learning, Deep learning, Cognitive science, Formal concept analysis, Brain modeling, Uncertainty analysis, Common sense reasoning, Natural language processing, Natural computing
Subtrack 3.5 Neural Network:
Artificial neural networks, Dimensionality reduction and manifold learning, Classification and clustering, Representation and analysis in pixel/voxel images, Support vector machines and kernel methods, Transfer learning, Semi-supervised learning and spectral methods, Model selection, Reinforcement learning and temporal models
Subtrack 3.6 Knowledge Discovery in Database:
Data mining & knowledge discovery, Image mining, Text mining, Computing with words, Web intelligence & semantic web, Social media mining, Crowd sourced computing
Subtrack 3.7 Cryptography:
Cryptography, Cryptology, Crypt-arithmetic, Visual cryptography, Steganography and Digital water marking

Track 4: Digital Signal and Image Processing

Subtrack 4.1:
Signal, image and video processing

Subtrack 4.2:
Audio and acoustic processing and analysis
Subtrack 4.3:
Multimedia analysis, indexing and retrieval
Subtrack 4.4:
Sensor array & multichannel signal processing
Subtrack 4.5:
Segmentation, features and descriptors
Subtrack 4.6:
Texture and color analysis
Subtrack 4.7:
Enhancement, restoration and filtering
Subtrack 4.8:
Image and video analysis and understanding
Subtrack 4.9:
Automatic speech and speaker recognition

Track 5: Modern Instrumentation, Control, Robotics and Automation

Subtrack 5.1 Networked and Distributed Control:
Intelligent control, Real time supervisory control, Adaptive control systems, Mobile autonomous systems
Subtrack 5.2 System Integration:
Embedded systems, Manufacturing systems, Sensors, Actuators, Modeling and simulation, MEMS and NEMS
Subtrack 5.3 Control Systems:
Real-time control, Intelligent control, Monitoring and supervision, Observers, Estimation and identification, Machine learning and pattern recognition, Nonlinear control, Robust control, Adaptive control, Optimal control, Digital control, Distributed and networked control, Control applications
Subtrack 5.4 Robotics:
Agriculture and field robotics, Robotic service and security, Entertainment and rescue robotics, Novel robotic locomotion
Subtrack 5.5 Mutliagent and Collaborative Systems:
MACS, Biorbotics, Biomechatronics, Amphibionics, Aerobatics-UMV, Telerobotics.
Subtrack 5.6 Mechatronics:
Mechatronics applications for control and automation of renewable energy, Mechatronics education, Virtual labs, e-learning in mechatronics, Bio-mechatronics, Autotronics, Mechanism design and applications

Track 6: Civil Engineering and Structural Design

Subtrack 6.1 Civil and Urban Engineering:
Structural & construction engineering, Road and bridge engineering, Geotechnical engineering, Hydraulic engineering, Costal engineering, Earthquake engineering, Materials engineering
Subtrack 6.2 Architecture and Design:
Urban planning and design, Building technology science, Art design and Landscape architecture, Aesthetic and landscape energy, Ecological control and intelligent control, Sustainable infrastructure
Subtrack 6.3 Architecture and Design:
Urban planning and design, Building technology science, Art design and Landscape architecture, Aesthetic and landscape energy, Ecological control and intelligent control, Sustainable infrastructure
Subtrack 6.4 Environmental Engineering:
Environment and climate change—Global warming, Green construction and environment protection, Environment science and health, Environment and social development, Environment and economic restructuring, Environment and pollution cost, Water supply and sewage processing
Subtrack 6.5 Transportation Engineering:
Acquisition, Processing and publishing of traffic information, Traffic guidance and forecast, Urban traffic control and congestion pricing, Vehicle safety and emission, Automatic incident detection, Intelligent transportation and logistics

Track 7: Real-time and Embedded Systems, Communication and Devices

Subtrack 7.1 Real-time Systems:
RTOS, RT scheduling, Middleware Systems, Media processing and transmission, RT aspects of wireless sensor network
Subtrack 7.2 Embedded Systems:
Embedded systems architecture, Mutli-core embedded systems—Dual core/icore, Embedded software and compilers, Ubiquitous and distributed ES, Reconfigurable computing architecture- Android
Subtrack 7.3 Cloud Computing:
X as a Service, Big data management and analysis, Virtualization technology, Cloud security and privacy, Cloud programming model and tools, Service oriented architecture
Subtrack 7.4 Embedded Systems and its Application in Healthcare:
DSP based medical devices and implants, Body sensor network, Wearable sensors and systems, Fuzzy systems and signals in medical electronics, Medical electronics

miniaturization and design, Smart materials and medical electronics, Artificial support for rehabilitation
Subtrack 7.5 Modern Communication Systems

Track 8: Advanced Optimization Techniques

Subtrack 8.1 Multiobjective Optimization:
Multiobjective programming (MOP), Multi Criterion Decision Making (MCDM), Minsum/Priority based goal programming
Subtrack 8.2 Hierarchical Decision Making:
Bi-Level Programming (BLP), Multi-Level Programming (MLP)
Subtrack 8.3 Decision Making in Imprecise environment:
Stochastic programming—CCP and DCP, Fuzzy programming- Fuzzy Decision, Fuzzy multiobjective programming, Fuzzy goal programming, Interval programming
Subtrack 8.4 Evolutionary Computation:
Evolutionary strategy, Evolutionary algorithm, Genetic algorithm, Multiobjective genetic algorithm (MOGA), Genetic optimization

Track 9: Biotechnology, Biomedical Instrumentation and Bioinformatics

Subtrack 9.1 Biotechnology:
Pharmaceutical biotechnology, Therapeutic drug, Nonobiotechnology, Plant biotechnology: cross breading, green revolution, Animal biotechnology: cloning and genetic engineering, Boitech health care: transgenic animals and clinical therapy, Biopharmaceuticals, Inoculation and incubation, Molecular modeling, Genome analysis
Subtrack 9.2 Biomedical Image Analysis:
Medical image mining, Vascular imaging, Structural biology, Computer aided detection and diagnosis, Medical imaging and diagnosis
Subtrack 9.3 Biomedical Instrumentation and Measurement:
Bio-sensors and bio MEMs, Medical robotics, Biosystem modeling, Biotelemetry and telemedicine, Brain Computer Interface (BCI)
Subtrack 9.4 Bioinformatics:
Computational biology, Genomics and proteomics, Sequence analysis, Structural bioinformatics, Immuno- and chemo-informatics, Transcriptomics, Next generation sequencing

Track 10: Outcome Based Education

Subtrack 10.1:
Preparing students for governance and leadership
Subtrack 10.2:
Industry academy partnerships
Subtrack 10.3:
Entrepreneurship and innovation
Subtrack 10.4:
Formulation of outcome based curriculum and syllabus
Subtrack 10.5:
Innovative methods in teaching and learning
Subtrack 10.6:
On-line engineering education
Subtrack 10.7:
Women in engineering
Subtrack 10.8:
Student voice in the transformation process
Subtrack 10.9:
Stakeholders feedback and satisfaction survey

Kalyani, India — Swapan Bhattacharyya
Kalyani, India — Sabyasachi Sen
Kalyani, India — Meghamala Dutta
Kalyani, India — Papun Biswas
Kolkata, India — Himadri Chattopadhyay

Conference Committee

Chief Patron
Sardar Jodh Singh, Chairman JIS Group, India

Patrons
Mr. Taranjit Singh, MD, JIS Group
Mr. S.S. Dattagupta, Director, JIS Group
Dr. Malay R. Dave, Principal, JIS College of Engineering

Mentors
Prof. Asit Guha, Advisor, JIS Group
Mr. Simarpreet Singh, Director, JIS Group
Mr. U.S. Mukherjee, Dy. Director, JIS Group
Mrs. Jaspreet Kaur, Trustee, JIS Group

Convener
Dr. Somsubhra Gupta, Dean—Academic Affairs, Department of Computer Science and Engineering and Information Technology

Co-convener(s)
Dr. P.K. Bardhan, Department of Mechanical Engineering
Dr. Sandip Bag, Department of Biomedical Engineering

International and Industry Relations

Chair
Mrs. Sila Singh Ghosh, Registrar, JISCE & VP (CR), JIS Group

Coordinator(s)
Dr. Mahuya Das, Department of Chemistry
Dr. Mrinal Kanti Das, Department of Business Administration

Organizing Chair(s)
Dr. Sabyasachi Sen, Dean—Research and Development, Department of Physics/Nano Science and Technology

Programme Chair(s)
Dr. Meghamala Datta, Department of Biomedical Engineering

Publication Chair
Dr. S.K. Mitra, Department of Electrical Engineering

Co-chair
Dr. Papun Biswas, Department of Electrical Engineering

Coordinator(s)
Dr. Indranath Sarkar, Department of Electronics and Communication Engineering
Mr. Subhajit Basu Chowdhury, Center of Management Studies
Mr. Partha Roy, Department of Electrical Engineering

Finance Chair
Dr. Anindya Guha, Department of Humanities

Coordinator(s)
Mr. Santanu Mondal, Department of Mathematics

Registration and Reception

Chair
Dr. Manas Paul, Department of Computer Application

Co-chair
Mrs. Pranati Rakshit, Department of Computer Science and Engineering

Coordinator(s)
Mrs. Suparna Dasgupta, Department of Information Technology
Mr. Rupak Bhattacherjee, Department of Mathematics

Hospitality (Travel, Accommodation, Refreshment)

Chair
Mr. Debashis Sanki, Department of Information Technology

Co-chair
Mr. Rajdeep Chowdhury, Department of Computer Application

Coordinator(s)
Mr. Rakesh Naskar, Department of Electrical Engineering
Mr. Debashis Majumdar, Department of Mathematics

Venue Preparation and Stage Management

Chair
Dr. Karabi Ganguly, Department of Biomedical Engineering

Coordinator(s)
Mr. Bikramjit Paul, Department of Computer Application
Dr. Subhamoy Singha Roy, Department of Physics

Communication (Website, E-Communication Publicity and Media Outreach)

Chair
Dr. Swapan Bhattacharya, Department of Electronics and Communication Engineering

Co-chair
Mr. Saumyabrata Saha, Department of Information Technology

Coordinator(s)
Dr. Biswarup Neogi, Department of Electronics and Communication Engineering
Mr. Bikash Dey, Department of Electronics and Communication Engineering

Volunteer and Student Management

Chair
Mr. Apurba Pual, Department of Computer Science and Engineering

Co-chair
Mrs. Ranjana Ray, Department of Electronics and Communication Engineering

Coordinator(s)
Dr. Rimi Ghosh, Department of Electronics and Instrumentation Engineering
Mr. Aniruddha Biswas, Department of Information Technology

Contents

Part I Nonconventional Energy and Advanced Power System

Comparison of Solar and Jovian Radio Emission at 20.1 MHz 3
Debojyoti Halder and Bipasa Raha

Characteristics of Solar PV Array Implemented in Matlab Software 11
Gourab Das, M. De, S. Mandal and K.K. Mandal

Determination of Yield Coefficients of Methane and Carbon Dioxide in Methanogenesis to Predict the Overall Performance of a Biogas Digester 21
Joyoti Biswas, Ranjana Chowdhury and Pinaki Bhattacharya

A Novel Algorithm for Economic Load Dispatch Using a New Optimization Technique 29
S. Mandal, G. Das, M. De, B. Tudu and K.K. Mandal

Investigating Economic Emission Dispatch Problem Using Improved Particle Swarm Optimization Technique 37
Meenakshi De, Gourab Das, S. Mandal and K.K. Mandal

Design of Rule-Based Load Frequency Controller for Multi-machine System 47
Jyotirmoy Biswas and Parthasarathi Bera

Soft Computing Approach to Electrical Transmission Network Congestion Management 59
Debapriya Sur Mukhopadhyay, Reshmi Chanda, Debjani Chakraborti and Papun Biswas

Effect of Surface-to-Volume Ratio on Eigenenergy in Quantum Ring 71
Swapan Bhattacharyya, Sourish Haldar and Arpan Deyasi

Part II Nanotechnology and Applications

Unusual Conductance Characteristics in Single Electron Transistor 81
Arkadeep Paul, Ritabrata Chakraborty, Arpan Deyasi and Shrabani Nayak

Conductivity, Dielectric and Impedance Property Analysis of $Y_{1-X}La_xCrO_3$ Nanoparticles 89
R. Sinha, S. Basu and A.K. Meikap

Part III Pattern Recognition and Machine Intelligence

Generic Object Detection Framework with Spatially Pooled Features 99
K. Venkatachalapathy, K. Kishore Anthuvan Sahayaraj and V. Ohmprakash

Hybrid PSACGA Algorithm for Job Scheduling to Minimize Makespan in Heterogeneous Grids 107
Amit Chhabra and Oshin

Survey of Various Real and Non-real-time Scheduling Algorithms in Mobile Ad hoc Networks 121
Abu Sufian, Anuradha Banerjee and Paramartha Dutta

Double Ended Bottom-Up Approach of Data Warehouse Architecture Employing Country ISO Code Engendered Cryptographic Algorithm 135
Rajdeep Chowdhury, Ipshita Bhattacharya, Nirmita De and Subhajit Saha

Factors Affecting Crime Against Women Using Regression and K-Means Clustering Techniques 149
Bhajneet Kaur, Laxmi Ahuja and Vinay Kumar

Energy Efficient Data Gathering in Wireless Sensor Networks Using Rough Fuzzy C-Means and ACO 163
Sanjoy Mondal, Saurav Ghosh and Pratik Dutta

AI Doctor: An Intelligent Approach for Medical Diagnosis 173
Sumit Das, S. Biswas, Aditi Paul and Aritra Dey

An Online Trend Detection Strategy for Twitter Using Mann–Kendall Non-parametric Test 185
Sourav Malakar, Saptarsi Goswami and Amlan Chakrabarti

Part IV Digital Signal and Image Processing

Full Reference Image Quality Assessment: A Survey 197
Baisakhi Sur Phadikar, Goutam Kumar Maity and Amit Phadikar

A Framework for Face Recognition Based on Fuzzy Minimal Structure Oscillation and Independent Component Analysis 209
Sharmistha Bhattacharya (Halder) and Srijita Barman Roy

Part V Modern Instrumentation, Control, Robotics and Automation

Computing Reflectance of Three-Layer Surface Plasmon-Based Sensor at Visible Spectra 221
Pratibha Verma, Arpan Deyasi and Pratiti Paul

Newton–Krylov Subspace Method to Study Structure Parameter Optimization in Rib Waveguide Communication 229
Sucharita Bhattacharyya and Anup Kumar Thander

Fuzzy-Tuned SIMC Controller for Level Control Loop 239
Ujjwal Manikya Nath, Chanchal Dey and Rajani K. Mudi

Utilization of Electromagnetic Sensor for Structural Characterization of Steels During Processing and in-Service Components 247
Rajat K. Roy, M. Premkumar, Ashis K. Panda and Amitava Mitra

Developing a Smart Navigator for Surveillance in Unmanned Zones 255
Pooja Nag, Sumit Shinde and Kapil Sadani

A Simple Flow Measurement System for Rotameters Using Webcam 265
Pooja Nag, Sumit Shinde, Dayananda Nayak and Kapil Sadani

Sensor Search Using Clustering Technique in a Massive IoT Environment 271
Nandhakumar Ramachandran, Varalakshmi Perumal, Sakithya Gopinath and Monika Jothi

Assessment of Occupational Risks in Construction Sites Using Interval Type-2 Fuzzy Analytic Hierarchy Process 283
Joy Debnath and Animesh Biswas

Comparative Evaluation of Capacity Analysis and Probability Distribution of Elements for Different Iterative Values of MIMO 299
Sutanu Ghosh

Server Utilization-Based Smart Temperature Monitoring System for Cloud Data Center 309
Sudipta Sahana, Rajesh Bose and Debabrata Sarddar

Part VI Civil Engineering and Structural Design

Assessment of Uncontrolled Intersections Through Calibration of VISSIM for Indian Traffic Conditions 323
Suprabeet Datta

Part VII Real-time and Embedded Systems, Communication and Devices

Electromagnetic Band Structure Computation of Metamaterial/Air Composition from First Principle for Optical Filter Application 341
Bhaswati Das and Arpan Deyasi

Mobility Prediction for Dynamic Location Area in Cellular Network Using Hidden Markov Model 349
Nilesh B. Prajapati and D.R. Kathiriya

Neighbor Constraint Traffic Centric Distributed Sinkhole Detection and Mitigation Approach for Quality of Service Improvement in Wireless Sensor Networks 357
K. Devibala, S. Balamurali, A. Ayyasamy and M. Archana

Moving Object Detection Using Local Binary Pattern and Gaussian Background Model 367
A.P. Athira, Midhula Vijayan and R. Mohan

Design of CMOS Integrator Circuit for Sigma Delta ADC for Aerospace Application 377
Deepak Prasad and Vijay Nath

Aspects of Low-Power High-Speed CMOS VLSI Design: A Review 385
Prolay Ghosh, Tanusree Saha and Barsha Kumari

An Intelligent Beta Reputation and Dynamic Trust Model for Secure Communication in Wireless Networks 395
S. Sathish, A. Ayyasamy and M. Archana

Effect of Incidence Angle on Optical Bandwidth in Ternary Photonic Crystal for Filter Application 403
Romi Dey, Meenakshi Banerjee, Antara Das and Arpan Deyasi

Energy-Efficient Connected Target Coverage in Multi-hop Wireless Sensor Networks 411
Swagata Biswas, Ria Das and Punyasha Chatterjee

Optimal Sink Placement in Wireless Sensor Networks to Increase Network Performance 423
Mir Md. Sajid Sarwar and Punyasha Chatterjee

Performance Evaluation of Flash ADCs in Presence of Offsets Using Hot Code Generator and Bit Swap Logic (BSL) 435
Pranati Ghoshal and Sunit Kumar Sen

Part VIII Advanced Optimization Techniques

Optimal Sensing Time Analysis for Efficient Data Transmission in Multiple Amplify-Forward Cognitive Relay Assisted Network 449
Sutanu Ghosh, Aditya Chaudhuri and Sayantani Ghosh

R Implementation of Bayesian Decision Theoretic Rough Set Model for Attribute Reduction 459
Utpal Pal, Sharmistha Bhattacharya (Halder) and Kalyani Debnath

CBSTD: A Cloud Based Symbol Table Driven DNA Compression Algorithm 467
Annwesha Banerjee Majumder and Somsubhra Gupta

Optimization of E-Jet Based Micro-manufacturing Process Using Desirability Function Analysis 477
Raju Das, Amit Kumar Ball and Shibendu Shekhar Roy

A Heuristic Path Search for Congestion Control in WSN 485
Ganesan Sangeetha, Muthuswamy Vijayalakshmi, Sannasi Ganapathy and Arputharaj Kannan

Alignment-Independent Sequence Analysis Based on Interval Distribution: Application to Subtyping and Classification of Viral Sequences 497
Uddalak Mitra and Balaram Bhattacharyya

Part IX Biotechnology, Biomedical Instrumentation and Bioinformatics

Automated Glaucoma Detection from Fundus Images of Eye Using Statistical Feature Extraction Methods and Support Vector Machine Classification 511
Abhishek Dey and Kashi Nath Dey

Classification of EEG Analysis for Sleep Detection Using Wavelet Transform and Neural Network 523
G.K. Rajini, Shaik Naseera, Shikha Pandey and Akshaya Bhuvaneswaran

On Blood Flow Through an Overlapping Stenosed Artery 535
Anuprava Biswas

An Energy-Efficient Congestion Avoidance Priority-Based Routing Algorithm for Body Area Network 545
Annwesha Banerjee Majumder and Somsubhra Gupta

Dynamic Thresholding with Tabu Search for Detection of Hard Exudates in Retinal Image 553
Diptoneel Kayal and Sreeparna Banerjee

Application of Chaos Game in Tri-Nucleotide Representation for the Comparison of Coding Sequences of β-Globin Gene 561
Subhram Das, Nobhonil Roy Choudhury, D.N. Tibarewala and D.K. Bhattacharya

Computational Methodologies Followed in Structure Based In-Silico Drug Design: An Example 569
Indrani Sarkar and Sanjay Goswami

A Study on Data Acquisition and Analysis in Real Time for Prediction of Cardiovascular Disease in Real Time Using Indigenous Low-Cost CHMS 575
Meghamala Dutta, Deepneha Dutta, Swati Sikdar, Sourav Dutta and Shreya Dutta

Part X Outcome Based Education

Some Limitations of Outcome-Based Education 591
Avik Sanyal and Rajashree Gupta

Activity-Based Learning (ABL) for Engaging Engineering Students 601
Aniruddha Biswas, Sumit Das and Suchandan Ganguly

Realization of Outcome-Based Education through Teaching Learning Methodology in Engineering Institution 609
Rajdeep Chowdhury, Soumyabrata Saha, Sudipta Sahana and Debashis Sanki

Author Index 619

Editors and Contributors

About the Editors

Dr. Swapan Bhattacharyya (M'06) received his B.Sc. (Hons.) degree in Physics from the University of Calcutta in 1982, the Integrated M.Tech. (B.Tech & M.Tech) degree in Radio Physics and Electronics in 1987 and the Ph.D. degree, all from the University of Calcutta. In 1988, he joined the Tata Group of Companies (TGC) at Jamshedpur, India as a Senior Customer Support Engineer, where he became the Branch Manager (Customer Support) in 1998 and in the next year he joined the Corporate Office—Mumbai as a Corporate Manager.

In May 2004, he joined the Asansol Engineering College (India) as a System Manager, where he became the Head of the Department of Computer Science and Engineering since December 2004. At present, he is the Head of the Department of Electronics and Communication Engineering at JIS College of Engineering (An Autonomous Institution). He is continuing his research in the field of semiconductor nanostructures with special focus on quantum dots for information and communication applications. He has also special interest in natural language processing, and use of IT for rural development. Since 2013, he has been one of visiting faculties in Radio Physics and Electronics Department of Calcutta University. Dr. Bhattacharyya was a recipient of a National Scholarship from Government of India. He worked in academic administration of several reputed technical institutes and universities in India. He has published many technical research papers in international journals, in national and international conferences. He served as the convener of technical committee for several national conferences and seminars. He is also the reviewer of IEEE, Journal of Applied Physics. He is a member of IEEE, IEEE Photonics Society and life member of IE (Institute of Engineers, India).

Dr. Sabyasachi Sen received his Bachelor, Masters and Ph.D. degree from the prestigious University of Calcutta. He has been associated with academics for nearly 15 years and with research and development activities for 13 years. Over these years, he has served the institute at the various administrative position including HOD, Physics and Nanoscience and Technology; Dean R&D, Secretary Academic Council of Autonomous JIS College of Engineering. He has 21 publications in reputed journals which include premier journals like Physical Review Letters (Phys. Rev. Lett. 109, 257204 (2012); Impact factor: 7.5), Journal of American Chemical Society (J. Am. Chem. Soc. 132 (2010) 15334; Impact factor: 11, Research Highlight in Nature India, entitled 'Electron spin memory'), Journal of Physical Chemistry C (Impact factor: 4.8), Physical Review B (Impact factor: 3.69), Chem Phys Chem (Impact factor: 3.349), Physical Chem. Chem. Physics (Impact factor: 4.493) and many more. He has published one book chapter in the Hand of Book of Nanophysics. He is presently supervising three research projects funded by DST. In 2008, he received Sir Acharya Prafulla Chandra Roy Award (2007–08) from University of Calcutta for contribution in research activities. Apart from this he along with HOD, Chemistry of the JIS College of Engineering has initiated Masters' Program in Nanoscience and Technology, the technology of next generation.

Dr. Meghamala Dutta is presently with JIS College of Engineering as Examination Controller and Senior Faculty in the Department of Biomedical Engineering. Dr. Dutta was formerly the Head, Department of Biomedical Engineering and successfully completed 4 years. She is multifaceted leader, who has worked in diverse field across the academia in managing several cutting-edge research projects funded by Government of India, organizing national and international level conferences with participation from Harvard Medical School and other organizations of repute, authoring several papers in national and international journals including a book chapter. She has over 12 years of experience in industry and academia and has played a key role in technical course design, implementation of new courses, identification of thrust areas essential for industry academia interfacing, governance and research. A member of IEEE, Dr. Dutta is also the Nominee Chairman under MHRD run Kendriya Vidyalaya. Dr. Dutta has been associated with several industries in developing technical projects, in undergraduate students beneficial for placement. She has been the Nodal Officer for EAP (Equity Action Plan) the prestigious World Bank funded TEQIP Project and is involved with establishing finishing schools, soft skills and communication skills development training programs.

Dr. Papun Biswas received his M.Tech and Ph.D. degree from the University of Calcutta and University of Kalyani, respectively. Dr. Biswas is presently the Teacher in-Charge of the Department of Electrical Engineering, JIS College of Engineering. He has been associated with academics for nearly 12 years and with research and development activities for 9 years. He has 28 publications in international journals, book chapters and international conference proceedings which include Elsevier, Springer and IEEE. He is also the reviewer of IEEE Transactions on Power Systems, International Journal of Electrical Power and Energy Systems, Applied Mathematics and Computation, European Journal of Operational Research. He is a member of IEEE and IEEE CSI Society. He has been the Nodal Officer for Procurement of the prestigious World Bank funded TEQIP Project. He is continuing his research on application of soft computing techniques in the field of electrical engineering.

Prof. Himadri Chattopadhyay is currently working as Associate Professor, Department of Mechanical Engineering, Jadavpur University, Kolkata, West Bengal, India. He received his Ph.D. degree in the year of 2002. Professor Chattopadhyay has more than 70 publications in various reputed international journals and international conferences. He has received fellowship and reinvention fellowship from DAAD in the year of 1997 and 2003, respectively. His areas of research interest include transport phenomena/convection, turbulence, heat transfer, augmentation, biofluid dynamics and material processing.

Contributors

Laxmi Ahuja AIIT, Amity University, Noida, UP, India

M. Archana Faculty of Engineering and Technology, Department of Computer Science and Engineering, Annamalai University, Chidambaram, Tamil Nadu, India

A.P. Athira Department of CSE, NIT Trichy, Trichy, Tamil Nadu, India

A. Ayyasamy Faculty of Engineering and Technology, Department of Computer Science and Engineering, Annamalai University, Chidambaram, Tamil Nadu, India

S. Balamurali Department of Computer Application, Kalasalingam University, Virudhunagar, Tamil Nadu, India

Amit Kumar Ball Department of Mechanical Engineering, National Institute of Technology, Durgapur, Durgapur, India

Anuradha Banerjee Department of Computer Application, Kalyani Government Engineering College, Kalyani, WB, India

Meenakshi Banerjee Department of Electronics and Communication Engineering, RCC Institute of Information Technology, Kolkata, India

Sreeparna Banerjee Maulana Abul Kalam Azad University of Technology, Kolkata, West Bengal, India

S. Basu Department of Physics, NIT Durgapur, Durgapur, West Bengal, India

Parthasarathi Bera Department of Electrical Engineering, Kalyani Government Engineering College, Kalyani, India

D.K. Bhattacharya Bio-Science & Engineering, Jadavpur University, Kolkata, India

Ipshita Bhattacharya Department of Electronics and Communication Engineering, JIS College of Engineering, Kalyani, Nadia, West Bengal, India

Pinaki Bhattacharya Department of Chemical Engineering, Jadavpur University, Kolkata, India

Sharmistha Bhattacharya (Halder) Department of Mathematics, Tripura University (A Central University), Agartala, Tripura, India

Balaram Bhattacharyya Department of Computer and System Sciences, Visva-Bharati University, Santiniketan, India

Sucharita Bhattacharyya Department of Applied Science and Humanities, Guru Nanak Institute of Technology, Kolkata, India

Swapan Bhattacharyya Department of Electronics and Communication Engineering, JIS College of Engineering, Kalyani, India

Akshaya Bhuvaneswaran School of Electrical Engineering, VIT University, Vellore, India

Animesh Biswas Department of Mathematics, University of Kalyani, Kalyani, India

Aniruddha Biswas Information Technology, JIS College of Engineering, Kalyani, Nadia, West Bengal, India

Anuprava Biswas Department of Mathematics, Krishnath College Berhampore, Murshidabad, West Bengal, India

Joyoti Biswas Department of Chemistry, Dr. Sudhir Chandra Sur Degree Engineering College, Kolkata, India

Jyotirmoy Biswas Department of Electrical Engineering, Kalyani Government Engineering College, Kalyani, India

Papun Biswas JIS College of Engineering, Kalyani, WB, India

S. Biswas DETS, University of Kalyani, Kalyani, India

Swagata Biswas School of Mobile Computing and Communication, Jadavpur University, Kolkata, India

Rajesh Bose Department of Computer Science and Engineering, University of Kalyani, Kalyani, Nadia, West Bengal, India

Amlan Chakrabarti Kolkata, India

Debjani Chakraborti Narula Institute of Technology, Agarpara, WB, India

Ritabrata Chakraborty Department of Electronics and Communication Engineering, RCC Institute of Information Technology, Kolkata, India

Reshmi Chanda Abacus Institute of Engineering and Management, Mogra, WB, India

Punyasha Chatterjee School of Mobile Computing and Communication, Jadavpur University, Kolkata, India

Aditya Chaudhuri Dr. Sudhir Chandra Sur Degree Engineering College, Kolkata, India

Amit Chhabra Department of Computer Engineering and Technology, Guru Nanak Dev University, Amritsar, India

Nobhonil Roy Choudhury Computer Science & Engineering, Narula Institute of Technology, Kolkata, India

Rajdeep Chowdhury Department of Computer Application, JIS College of Engineering, Kalyani, Nadia, West Bengal, India

Ranjana Chowdhury Department of Chemical Engineering, Jadavpur University, Kolkata, India

Antara Das Department of Electronics and Communication Engineering, RCC Institute of Information Technology, Kolkata, India

Bhaswati Das Department of Electronics and Communication Engineering, RCC Institute of Information Technology, Kolkata, India

Gourab Das Department of Power Engineering, Jadavpur University, Kolkata, India

Raju Das Department of Mechanical Engineering, National Institute of Technology, Durgapur, Durgapur, India

Ria Das School of Mobile Computing and Communication, Jadavpur University, Kolkata, India

Subhram Das Computer Science & Engineering, Narula Institute of Technology, Kolkata, India

Sumit Das Information Technology, JIS College of Engineering, Kalyani, Nadia, West Bengal, India

Suprabeet Datta Department of Civil Engineering, JIS College of Engineering, Kalyani, Nadia, West Bengal, India

Meenakshi De Department of Power Engineering, Jadavpur University, Kolkata, India

Nirmita De Department of Electronics and Communication Engineering, JIS College of Engineering, Kalyani, Nadia, West Bengal, India

Joy Debnath Department of Mathematics, University of Kalyani, Kalyani, India

Kalyani Debnath Department of Mathematics, Tripura University (A Central University), Agartala, Tripura, India

K. Devibala Department of Computer Science, Ayya Nadar Janaki Ammal College, Sivakasi, Tamil Nadu, India

Abhishek Dey Department of Computer Science, Bethune College, Kolkata, India

Aritra Dey Electrical Engineering, JIS College of Engineering, Kalyani, India

Chanchal Dey Department of Applied Physics Instrumentation and Control Engineering, University of Calcutta, Kolkata, India

Kashi Nath Dey Department of Computer Science and Engineering, University of Calcutta, Kolkata, India

Romi Dey Department of Electronics and Communication Engineering, RCC Institute of Information Technology, Kolkata, India

Arpan Deyasi Department of Electronics and Communication Engineering, RCC Institute of Information Technology, Kolkata, India

Deepneha Dutta Department of Biomedical Engineering, JIS College of Engineering, Kalyani, India

Meghamala Dutta Department of Biomedical Engineering, JIS College of Engineering, Kalyani, India

Paramartha Dutta Department of Computer & System Science, Visva-Bharati University, Santiniketan, WB, India

Pratik Dutta Department of Computer Science and Engineering, University of Calcutta, Kolkata, India

Shreya Dutta IIMSAR & Dr. BCRHH, Haldia, India

Sourav Dutta IBM India Ltd., New Delhi, India

Sannasi Ganapathy School of Computing Science and Engineering, VIT University-Chennai Campus, Chennai, India

Suchandan Ganguly Information Technology, JIS College of Engineering, Kalyani, Nadia, West Bengal, India

Prolay Ghosh Department of Information Technology, JIS College of Engineering, Kalyani, India

Saurav Ghosh A.K. Choudhury School of I.T., University of Calcutta, Kolkata, India

Sayantani Ghosh Jadavpur University, Kolkata, India

Sutanu Ghosh Indian Institute of Engineering Science and Technology, Shibpur, Howrah, India

Pranati Ghoshal Techno India, Kolkata, India

Sakithya Gopinath Department of Computer Technology, MIT Campus, Anna University, Chennai, India

Sanjay Goswami Department of Computer Applications, Narula Institute of Technology, Kolkata, West Bengal, India

Saptarsi Goswami Kolkata, India

Rajashree Gupta Shantipur B.Ed College, Shantipur, Nadia, India

Somsubhra Gupta JIS College of Engineering, Kalyani, West Bengal, India

Sourish Haldar Department of Electronics and Communication Engineering, NIT Durgapur, Durgapur, India

Debojyoti Halder Department of Physics, Rishi Bankim Chandra Evening College, Naihati, West Bengal, India

Monika Jothi Department of Computer Technology, MIT Campus, Anna University, Chennai, India

Arputharaj Kannan Department of Information Science & Technology, Anna University, Chennai, India

D.R. Kathiriya Computer Center, Anand Agriculture University, Anand, Gujarat, India

Bhajneet Kaur AIIT, Amity University, Noida, UP, India

Diptoneel Kayal Maulana Abul Kalam Azad University of Technology, Kolkata, West Bengal, India

K. Kishore Anthuvan Sahayaraj Annamalai University, Chidambaram, India

Vinay Kumar Vivekananda Institute of Professional Studies, GGSIPU, New Delhi, India

Barsha Kumari Department of Information Technology, JIS College of Engineering, Kalyani, India

Goutam Kumar Maity Department of ECE, Netaji Subhash Engineering College, Garia, India

Annwesha Banerjee Majumder JIS College of Engineering, Kalyani, West Bengal, India

Sourav Malakar Kolkata, India

K.K. Mandal Department of Power Engineering, Jadavpur University, Kolkata, India

S. Mandal Department of Electrical Engineering, Jadavpur University, Kolkata, India

A.K. Meikap Department of Physics, NIT Durgapur, Durgapur, West Bengal, India

Amitava Mitra NDE & Magnetic Materials Gr., MST Division, CSIR-National Metallurgical Laboratory, Jamshedpur, India

Uddalak Mitra Department of Computer and System Sciences, Visva-Bharati University, Santiniketan, India

R. Mohan Department of CSE, NIT Trichy, Trichy, Tamil Nadu, India

Sanjoy Mondal A.K. Choudhury School of I.T., University of Calcutta, Kolkata, India

Rajani K. Mudi Department of Instrumentation and Electronics Engineering, Jadavpur University, Kolkata, India

Debapriya Sur Mukhopadhyay Chandernagore Municipal Corporation, Chandannagar, WB, India

Pooja Nag Department of Mechatronics, Manipal Institute of Technology, Manipal University, Manipal, Karnataka, India

Shaik Naseera School of Computing Science and Engineering, VIT University, Vellore, India

Ujjwal Manikya Nath Department of Instrumentation and Electronics Engineering, Jadavpur University, Kolkata, India

Vijay Nath VLSI Design Group, Department of ECE, Birla Institute of Technology, Ranchi, India

Dayananda Nayak Department of Instrumentation and Control, Manipal Institute of Technology, Manipal University, Manipal, India

Shrabani Nayak Department of Electronics and Communication Engineering, RCC Institute of Information Technology, Kolkata, India

V. Ohmprakash Annamalai University, Chidambaram, India

Oshin Department of Computer Engineering and Technology, Guru Nanak Dev University, Amritsar, India

Utpal Pal Department of Mathematics, Tripura University (A Central University), Agartala, Tripura, India

Ashis K. Panda NDE & Magnetic Materials Gr., MST Division, CSIR-National Metallurgical Laboratory, Jamshedpur, India

Shikha Pandey School of Electrical Engineering, VIT University, Vellore, India

Aditi Paul Electrical Engineering, JIS College of Engineering, Kalyani, India

Arkadeep Paul Department of Electronics and Communication Engineering, RCC Institute of Information Technology, Kolkata, India

Pratiti Paul Department of Electronics and Communication Engineering, NIT Agartala, Jirania, India

Varalakshmi Perumal Department of Computer Technology, MIT Campus, Anna University, Chennai, India

Amit Phadikar Department of Information Technology, MCKV Institute of Engineering, Liluah, India

Baisakhi Sur Phadikar Department of Computer Science & Engineering, MCKV Institute of Engineering, Liluah, India

Nilesh B. Prajapati Computer IT Engineering, B.V.M. Engineering College, Gujarat Technological University, Anand, India

Deepak Prasad VLSI Design Group, Department of ECE, Birla Institute of Technology, Ranchi, India

M. Premkumar NDE & Magnetic Materials Gr., MST Division, CSIR-National Metallurgical Laboratory, Jamshedpur 831007, India

Bipasa Raha Department of Physics, Techno India University, Kolkata, West Bengal, India

G.K. Rajini School of Electrical Engineering, VIT University, Vellore, India

Nandhakumar Ramachandran Department of Computer Technology, MIT Campus, Anna University, Chennai, India

Rajat K. Roy NDE & Magnetic Materials Gr., MST Division, CSIR-National Metallurgical Laboratory, Jamshedpur, India

Shibendu Shekhar Roy Department of Mechanical Engineering, National Institute of Technology, Durgapur, Durgapur, India

Srijita Barman Roy Department of Mathematics, Tripura University, Agartala, India

Kapil Sadani Department of Instrumentation and Control, Manipal Institute of Technology, Manipal University, Manipal, Karnataka, India

Soumyabrata Saha Department of Information Technology, JIS College of Engineering, Kalyani, Nadia, West Bengal, India

Subhajit Saha Department of Electronics and Communication Engineering, JIS College of Engineering, Kalyani, Nadia, West Bengal, India

Tanusree Saha Department of Information Technology, JIS College of Engineering, Kalyani, India

Sudipta Sahana Department of Computer Science and Engineering, JIS College of Engineering, Kalyani, Nadia, West Bengal, India

Mir Md. Sajid Sarwar School of Mobile Computing and Communication, Jadavpur University, Kolkata, India

Ganesan Sangeetha Department of Information Science & Technology, Anna University, Chennai, India

Debashis Sanki Department of Information Technology, JIS College of Engineering, Kalyani, Nadia, West Bengal, India

Avik Sanyal Centre for Management Studies, JIS College of Engineering, Kalyani, Nadia, India

Debabrata Sarddar Department of Computer Science and Engineering, University of Kalyani, Kalyani, Nadia, West Bengal, India

Indrani Sarkar Department of Physics, Narula Institute of Technology, Kolkata, West Bengal, India

S. Sathish Faculty of Engineering and Technology, Department of Computer Science and Engineering, Annamalai University, Annamalai Nagar, Tamil Nadu, India

Sunit Kumar Sen University of Calcutta, Kolkata, India

Sumit Shinde Department of Instrumentation and Control, Manipal Institute of Technology, Manipal University, Manipal, Karnataka, India; Department of Embedded Systems Engineering, Albert Ludwig University of Freiburg, Freiburg im Breisgau, Germany

Swati Sikdar Department of Biomedical Engineering, JIS College of Engineering, Kalyani, India

R. Sinha Department of Physics, Asansol Engineering College, Asansol, West Bengal, India; Department of Physics, NIT Durgapur, Durgapur, West Bengal, India

Abu Sufian Department of Computer Science, University of Gour Banga, Malda, WB, India

Anup Kumar Thander Department of Applied Science and Humanities, Guru Nanak Institute of Technology, Kolkata, India

D.N. Tibarewala Bio-Science & Engineering, Jadavpur University, Kolkata, India

B. Tudu Department of Power Engineering, Jadavpur University, Kolkata, India

K. Venkatachalapathy Annamalai University, Chidambaram, India

Pratibha Verma Department of Electronics and Communication Engineering, NIT Agartala, Jirania, India

Muthuswamy Vijayalakshmi Department of Information Science & Technology, Anna University, Chennai, India

Midhula Vijayan Department of CSE, NIT Trichy, Trichy, Tamil Nadu, India

About JIS College of Engineering

JIS College of Engineering was established in the millennium year 2000 by JIS Foundation and over the last decade grown in rapid strides to transform it into an Autonomous Institute. **JIS Group Educational Initiative** is the largest educational conglomerate in the state of West Bengal and leading private-sector educational group dedicated to impart demand-driven education in science, engineering, technology, management and medical science with highly laudable quality.

In the direction of thought of the honourable Chairman's dream—***Igniting Minds, Empowering Lives***—the institute has taken several new educational initiatives in the recent past. The prime focus continues to drive us towards achieving increased standard of technical education and wider research outreach programme. It is imperative, therefore, to be both creative and innovative in the approach to lead the institute towards a ***Centre of Excellence***.

The Institution is declared Autonomous by the University Grant Commission in 2011. The autonomous status conferred with qualification of 2(f) and 12(b). The Institution is approved by *All India Council for Technical Education (AICTE)* and affiliated to erstwhile *West Bengal University of Technology (WBUT) renamed MAKAUT*. The Institution was accredited by *National Assessment and Accreditation Council (NAAC)* in 2009 and qualified in Grade—'A' in its 2nd cycle in 2015. Its eligible technological departments are accredited by *National Board of*

Accreditation (NBA) or under reaccreditation process. Further, the Institution has qualified for the World Bank Grant under *Technical Education Quality Improvement Programme (TEQIP)* Phase II in the Subcomponent 1.2, for PG and Demand-driven Research & Development & Innovation, as the only private self-financed institution from the State of West Bengal.

Located in its sprawling 17.5 acres lush 'Green Campus' at Satellite Township of Kalyani, 55 km north of Kolkata, the college has intake of 3000+ students pursuing various degree engineering courses along with computer application and management programs. The campus is an ideal educational ambience with classrooms, laboratories, smart classrooms, seminar halls, conference halls, computer centre, central library, gymnasium, store and canteen with campus Wi-Fi connectivity.

The institution is moving ahead by allying, hand-holding and hand-shaking with leading universities and institutions of both the nation and the state. The institution is a learning-centric environment with authority always having a careful concern towards pedagogical aspects of the learning–teaching and making the faculty members up-to-date with high-end standard. Regular pedagogical training by IITs, technical training institute (viz. NITTTR) and by the corporate (viz. WIPRO Mission 10X) is a continuous process. The faculty members are trained in every available opportunity under NMEICT programme, ISTE workshops and are always encouraged to use global academic repositories like Massive Open Online Courses (MOOC) etc.

Being an engineering college, the institution is equally concerned for close networking and tie-ups with industry. Regular industrial events in mother-engineering as well as soft-engineering are phenomena. The setting up of "Collaborative Learning and Innovation Centre (CLIC)" with Tata Technologies in the recent past is a vibrant illustration. Collaboration with Texas Instruments or events like Infosys Campus Connect, IBM TGMC, Microsoft Faculty Connection, MSDN Academic Alliance, WIPRO Mission 10X Learning Approaches, and Oracle University from time to time are some illustrations. The AICTE funded Industry Institute Partnership Cell (IIPC) is recognition of the effort. Under autonomous framework, industry personnel have direct say in the curriculum design as Experts in the Board of Studies (BOS).

The institute has got a special bonding with national institutes especially with IITs. The overwhelming participation of the students in the ongoing initiative of National Project on Technology Enhanced Learning (NPTEL) in collaboration with IIT Kharagpur is worth mentioning. In Quality Circle, the institution has a set up of Quality Enhancement in Engineering Education (QEEE) frame-worked by MHRD and resourced by IIT Madras. The institution has an established Remote Centre (RC1335) of IIT Bombay and Resource Centre of IIT Kharagpur. The students could attend live classes from the teachers across the nation through Design Lab, Blended MOOC, Spoken Tutorial and Course Pack.

A numbers of value-added centers are set up recently to prepare the students in a role-ready/industry ready way. The Confederation of Indian Industry (CII) Innovation Centre, the British Council Library and Training Centre and collaboration initiative with BCC&I are illustrations.

Moreover, the students get opportunity to tune themselves by the Institutional Skill Development Programs. The IIPC and Finishing School have pro-active views regarding technical, communication and soft skills of the students. The students are provided with the programs like advanced (GATE coaching), remedial teaching to address the concerned of the advanced and weak students. Entrepreneurship has also been encouraged with eCell and AICTE sponsored Entrepreneurship Development Cell (EDC) in place.

The members of the Institution have enriched their experience and knowledge base through sharing and exchanging with all possible stake holders, viz. students, faculty, guardian, alumni, authority, employer and in extent with experts and regulatory bodies. This has facilitated our effort to improve institutional governance, networking and services to community and economy. The scrutiny, analysis, monitoring and evaluations of its programmes with community and self-appraisal are consistently practiced so as to identify the educational policies in a dynamic process.

The Research and Development is another strong aspect with faculty members receiving their Ph.D. award during their stay in the institution largely due to nurturing vision of the authority. Some more faculty members are on the avenues of getting their awards. Hosting and arranging R&D events like national and international conferences, workshops, symposiums and seminars are regular phenomena. In the recent past, the institution has hosted back-to-back international conferences, viz. IC3A2013 in the area computation and communication the proceedings of which have been published by the McGraw-Hill publication. This is followed by ICONCE 2014 in the area of Non conventional Energy technically cosponsored by IEEE regional section, the papers of which are available at IEEE Xplore. The experts, delegates and participants of event include stalwarts form USA, Germany, Japan, UK, the Netherlands and SAARC countries. The most recent ones ICAME 2015 in the area of medical electronics attended by the stalwarts from Harvard M.S the proceedings of which have been published by the Springer.

Part I
Nonconventional Energy and Advanced Power System

Comparison of Solar and Jovian Radio Emission at 20.1 MHz

Debojyoti Halder and Bipasa Raha

Abstract The Sun and the Jupiter both astronomical objects emit radiation in radio frequencies and have much importance in the Earth's atmosphere. We have recorded data of both Solar and Jovian signal on NASA's Radio Jove instrument at 20.1 MHz. The simultaneous visibility of the two objects—the Sun and the Jupiter—is of much importance in this study. In this paper, we have done the comparison using the Jove receiving system on those dates especially when the Jupiter rises at nighttime. It shows that the intensity of solar radio emission is well above Jovian emission intensity. It has been observed that solar radio bursts are stronger than Jovian bursts. They occur usually when there are sunspots on the visible corona while Jupiter noise storms are more likely to occur at nighttime and so they are much prominently recorded.

Keywords Solar radio emission · Jovian decametric emission · Radio antenna

1 Introduction

As the nearest star, the Sun is the brightest radiation source in most frequencies, down to the radio spectrum. Solar radio emission is one of the active phenomena of the sun where the stored magnetic energy is transformed into kinetic energy of highly accelerated particles via magnetic reconnection [1]. One of the characteristics of solar radiation is its time duration. Solar emissions observed at radio frequencies called bursts with duration less than 1 s are the main features of the basic energy release process in flares.

D. Halder (✉)
Department of Physics, Rishi Bankim Chandra Evening College,
Naihati 743165, West Bengal, India
e-mail: debojyoti829@gmail.com

B. Raha
Department of Physics, Techno India University, Sector-V, Salt Lake,
Kolkata 700091, West Bengal, India

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_1

Jupiter, on the other hand, has the largest planetary magnetosphere. Its giant magnetosphere acting both as trap and an accelerator of energetic charged particles produces intense belts of radiation. The accelerated charged particles emit radiation that depends on the energy of the particles [2]. The decametric radio waves have frequencies in the range between 8 and 40 MHz [3]. The present study has special importance from the point of view of radio astronomical observations on the Sun and the Jupiter over a tropical station in India.

2 Simultaneous Study of the Sun and the Jupiter

We have studied the Sun and the Jupiter in radio wavelengths mainly because of astronomical reasons. The Sun emits radiation in the whole range of E. M. spectrum, but Jupiter emits only in radio range. Radio waves are very long compared to rest of the electromagnetic spectrum. Some of the radio antennas can be utilized in any places where solar radio emission reaches. Again there are causes of studying the Jupiter along with the Sun. Jupiter emits radiation in 8–40 MHz frequency range while the Sun emits in 30–300 MHz radio range. The purpose of studying the Sun and the Jupiter simultaneously is that when the Jovian emission comes almost at the end, the solar emission starts. So when one goes from Jovian to solar emission or comes down from solar to Jovian emission there is a junction point in between. The study of this junction point is of much importance.

3 Recording Instrument for the Solar and Jovian Emission

The function of this antenna is to intercept the radio waves emitted by the Sun and the Jupiter and translate the signals into electrical current. The antenna is designed to receive signal corresponding to a frequency of 20.1 MHz as the Jupiter is mostly favorable around that frequency. Figure 1 shows the complete outline of Jove antenna and receiver setup.

The radio Jove system consists of three parts [3]:

(a) **The Radio Jove antenna**

It consists of two identical half-wave dipole antennas which can be phased together with a feed line. When converted to a full wave antenna, the gain is increased by 2 dB but the angle of maximum sensitivity reduces from 90° to 60°. It is particularly important that Jupiter be in the beam in order to realize the maximum antenna gain.

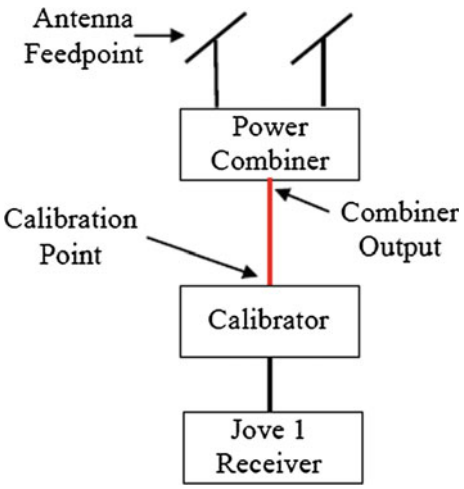


Fig. 1 Outline of the setup

(b) **The receiver**

The receiver is a simple direct conversion design which operates over a narrow band of frequencies centered at a 20.1 MHz. The receiver works by taking weak signal from the antenna and filtering out frequency. It converts the radio frequency to 3.5 kHz audio spectrum and amplifies the signal.

(c) **The calibrator**

The RF 2080 C/F calibrator contains both a calibrated 25000° noise source and a 20.1 MHz band pass filter. The band pass filter reduces the interference to the Jove receiver caused by strong international broadcasting stations and nearby power lines.

4 Comparison of Yearly Visibility of the Sun and the Jupiter and Dates for Jupiter Observation

Yearly simultaneous visibility of the two objects—the Sun and the Jupiter—is of much importance in this study. Figure 2 demonstrates the yearly visibility of the Sun and the Jupiter from 2010 to 2014 over the Kalyani sky. Visibility of the Jupiter is denoted by blue and that of the Sun is by yellow. The overlapping period is marked by green. According to this figure, the period of overlapping for these two objects changes for each year. We can see that in 2011 and 2012, the simultaneous visibility of the Sun and Jupiter was greater than the other years. The activities of these two objects were also higher than the adjacent years.

If the Jupiter is present in the daytime when the Sun is also present over our Kalyani sky, then there occurs much problem to distinguish the signals from these

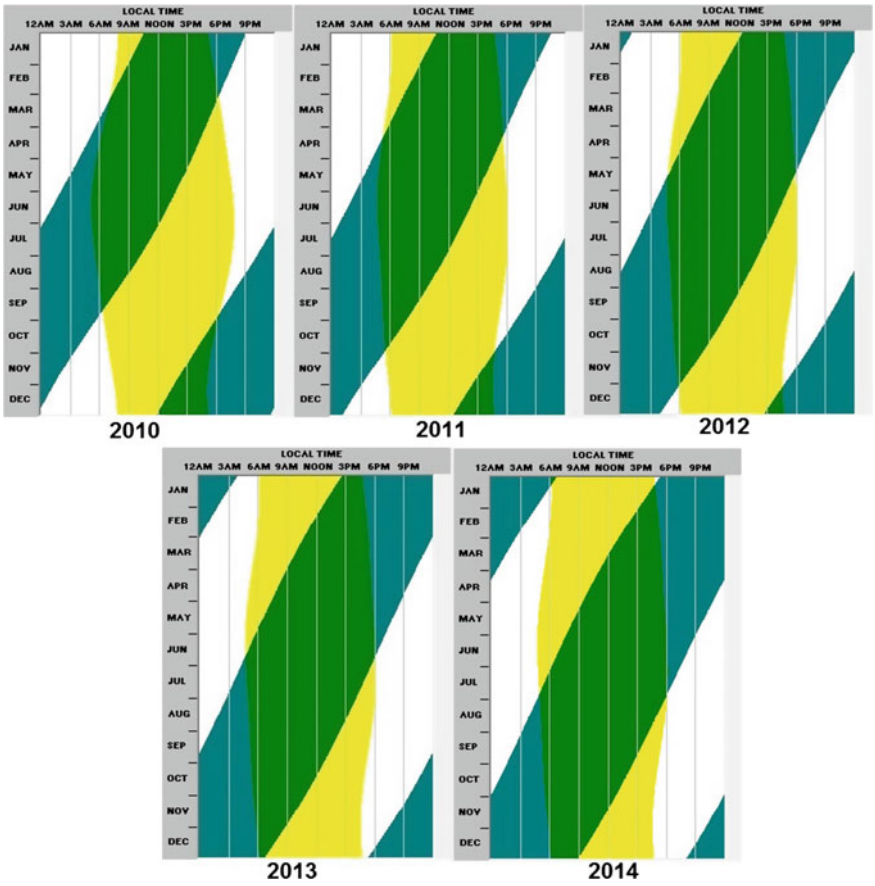


Fig. 2 Yearly visibilities of the Sun and the Jupiter from 2010 to 2014 over Kalyani

two objects. So we have chosen the dates very carefully while the Jupiter rises on the nighttime in our Kalyani sky so that we can observe both the objects in a 24 h day.

5 Comparison of Solar and Jovian Radio Bursts

We have done the comparison using our own Jove receiving system at 20.1 MHz. We have recorded both solar and Jovian radio data using the same instrument on these stipulated dates especially when the Jupiter rises with no Sun in Kalyani sky, i.e., at nighttime. Figure 3 demonstrates a 22 h plot of Solar and Jovian radio emission recorded on May 18, 2013. We can see the intensity of solar radio

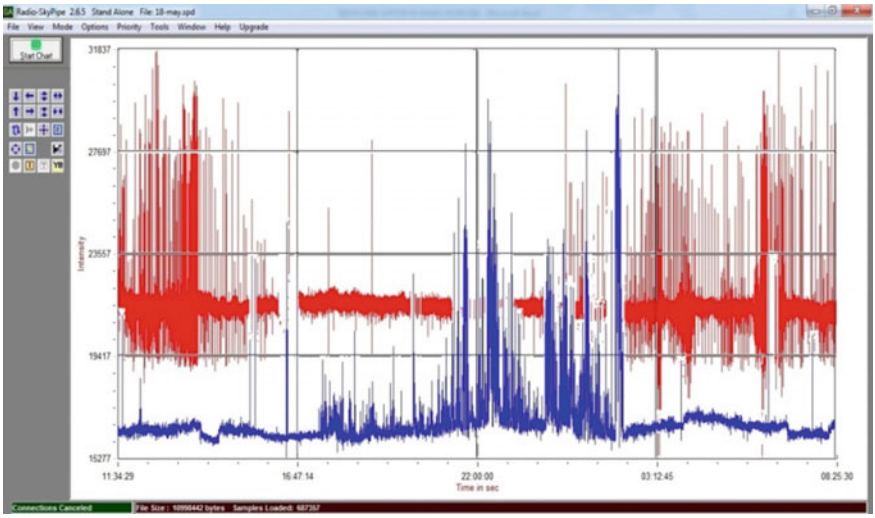


Fig. 3 Comparison of the solar and Jovian radio emissions

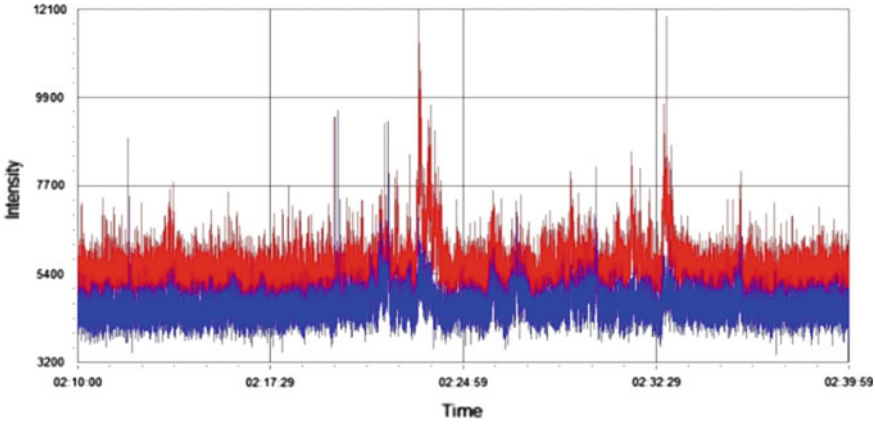


Fig. 4 Comparison of the solar and Jovian radio emissions on December 1, 2013

emission is well above Jovian emission intensity. The Jovian emission occurs at nighttime while solar emission occurs at daytime.

Figures 4 and 5 show the same as previous figure for comparing Solar and Jovian radio emissions. In Fig. 4 as we can see during daytime intensity level of solar radio signal always exceeds Jovian radio signal. But in Fig. 5 which shows a plot of morning time data intensity of Jovian signal exceeds solar signal in some time which is an interesting feature to be examined later.

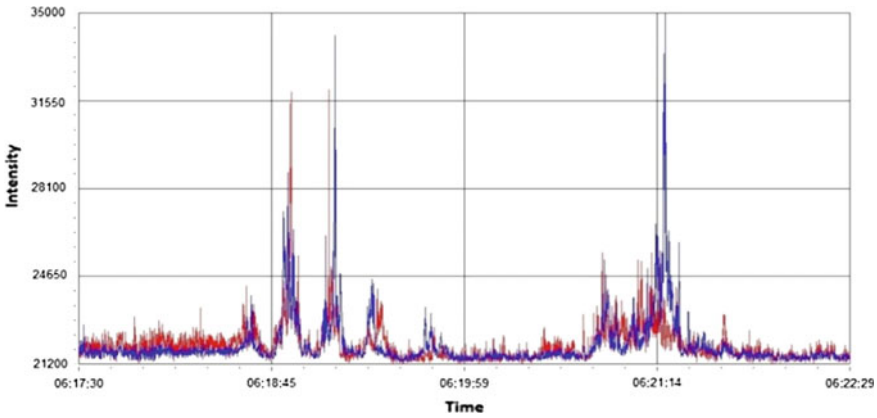


Fig. 5 Comparison of the solar and Jovian radio emissions on December 19, 2013

6 Discussion

Our observatory at the site of Kalyani is ideal due to reduced interference level. The radio emissions from Jupiter are highly variable which makes the possibility of detection above the galactic background. Solar radiation ionizes portions of Earth's atmosphere that can last from sunrise until several hours making the sunlit half of the upper atmosphere mostly opaque. Consequently, we have taken into account the observations of Jovian emissions after sunset and before sunrise. It has been observed that solar radio bursts are stronger than Jovian bursts. They occur usually when there are sunspots on the visible corona. Individual solar bursts generally last for about 30 s only but there is rapid onset followed by a slow decay. The resulting record on Skypipe looks like a shark fin. On the other hand, Jovian noise storms are more likely to occur at nighttime and so they are much prominently recorded. A radio noise storm of Jupiter can have duration of a few minutes to a couple of hours. Jupiter has two types of bursts—short duration bursts (S-bursts) that sound like popcorn popping while the other type is named as long duration bursts (L-bursts) sounds like ocean wave breaking up on a beach. Often we get a chance to listen combined activity of both L- and S-bursts. As we continue to gather more information on Radio Jove and also gathering electromagnetic waves from the Sun, we came to the conclusion that there should be further work on this subject. Experiments include listening for more solar bursts near in the future and also identifying how this emission from the Sun can affect the communication on Earth. By doing all of this, we could come and find a way in which our communication can be stronger and not be disturbed, nor interrupted by any solar or Jovian radio emission.

References

1. Culhane, J.L., Phillips, K.J.H.: Solar X-ray bursts at energies less than 10 keV observed with OSO-4. Solar Physics, vol. 11, p. 117. Springer Science-Business Media, Berlin (1970)
2. Khurana, K.K., Kivelson, M.G.: The configuration of Jupiter's magnetosphere. In: Bagenal, F., Dowling, T.E., McKinnon, W.B. (eds.) Jupiter: The Planet, Satellites and Magnetosphere. Cambridge University Press, Cambridge (2004)
3. Bhattacharya, A.B., Joarder, S., Nag, A., Halder, D., Debnath, M.: Reception of Jovian radio signals in a tropical station. Int. J. Eng. Sci. Technol. **2**, 5704–5713 (2010)

Characteristics of Solar PV Array Implemented in Matlab Software

Gourab Das, M. De, S. Mandal and K.K. Mandal

Abstract In recent time due to diminishing fossil fuel reserves and increasing social pressure in terms of environmental pollution, complex power system has no other option except to seek for another possibility of alternative energy to eliminate the environmental pollution. Solar energy accounts for the most of the renewable energy on earth. PV module is a particular procedure to accomplish the generation of electric power from solar radiation using semiconductor devices which exhibit photovoltaic effects. The concepts of a PV module and its characteristics have been studied. The P–V and V–I curves have been obtained at changing solar irradiation levels and temperatures.

Keywords Photovoltaic (PV) · Direct current (DC) · Current over voltage (V–I) · Power over voltage (P–V)

1 Introduction

The alarming rate of fossil fuel shortage, oil prices hike, global warming, and potential ill effects in environment, besides, gives the support by Government as well private agencies to develop environment friendly energy resources have put forward renewable power generation as an lucrative domain for research in this field. PV cell power is a sustainable energy technology that has seen extensive growth rate from the 1980s. Both the scaling down in cost and the progress in use have been substantial in solar power generation technology. Renewable energy systems and its subsequent implementation have nowadays become a suitable technical option to produce clean and affordable electric power. To produce a computer model of performance characteristics of different types of crystalline silicon photovoltaic

G. Das (✉) · M. De · K.K. Mandal
Department of Power Engineering, Jadavpur University, Kolkata 700098, India
e-mail: gourabdas.ju@gmail.com

S. Mandal
Department of Electrical Engineering, Jadavpur University, Kolkata 700032, India

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_2

modules/strings/arrays, novel theoretical approach has been the major focused area in recent literature [1]. In the recently developed model, the performance of the PV module at any solar irradiance and cell temperature can be simulated. The yield characteristic of solar PV module depends upon solar irradiation, cell temperature, and the output voltage [2]. An optimization-based approach for total cross-tied connected modules in a photovoltaic array is presented [3]. MATLAB coding-programmed modeling and simulation of PV module are presented by focusing on the effects of partial shading on the output of the photovoltaic (PV) systems [4]. Nowadays, PV array has been widely used in the electrical power system which is capable of generating DC electricity without environmental impact [5]. The PV module has been implemented by one-diode or two-diode models [6]. PV array has nonlinear characteristics and it is effective to design in maximum power point tracking. The PV module has semiconductor static device and there is no moving parts, so its maintenance cost is less. Simulation program of PV module has been implemented by series–parallel combination using some simulation parameter constant temperature and varying solar irradiance [7]. A solar module is modeled by attaching a number of cells in series to get a usable voltage which can be utilized in practice [8]. The C-language programming-based PV string simulation strategy was proposed [9] to simulate the effects of partial shading condition and develop new maximum power point tracking (new MPPT) algorithm for power conversion methods. The behavior of hybrid system hiring sustainable energy sources and changeable time is investigated [10] while providing a continuous steady supply. The work focused on modeling of a PV–wind hybrid system in Matlab/Simulink and the obtained model is useful for simulation of a grid-connected hybrid PV–wind system. The main aim of this research work is focused on the investigations on the modeling and simulation of solar energy sources.

1.1 Extraterrestrial Irradiation of Solar

In the core of the Sun originates innermost temperature 10^7 K by the nuclear fusion process and internal radiation flux has dissimilar spectral distribution. This inner radiation assimilates in the external submissive layers which are heated about 5800 K and it is source of radiation with a comparatively uninterrupted spectral distribution. The radiant flux varies throughout the year. So in this paper a different radiant flux (W/m^2) has been considered in Table 1.

1.2 Radiation Components of Solar Beam

Solar radiation fall on atmosphere from the direction of the Sun is the solar ethereal beam radiation. Underneath the atmosphere at the earth's uppermost layer, the radiation will be noticeable by the direction of the Sun disk (like a thin circular

Table 1 Simulation parameters for constant temperature and varying solar radiance

Simulation parameters	Values
Varying solar radiance intensities (S)	200, 400, 500, 800, 1000 W/m^2
Temperature of cell (T)	25 + 273
Reference temperature (Tref)	40 + 273
Short-circuit temperature coefficient (k_i)	0.0023 mA/°C
Boltzmann's constant (k)	1.38065×10^{-23}
Reverse saturation current (I_{rr})	21×10^{-10}A
Cell saturation current (Iscr)	0.75 mA
Area of the module (A)	0.40 m^2
Current temperature coefficient (α)	0.473 mA/°C
Voltage temperature coefficient (β)	636 V/°C
No. of cells connected in parallel (Np)	4
No. of cells connected in series (Ns)	100

object) in the direct beam and also other form of direction is diffuse radiation. The proportion between the beam radiance and the total radiance thus varies from about 0.9 on clear day to zero on complete foggy day (Figs. 1 and 2).

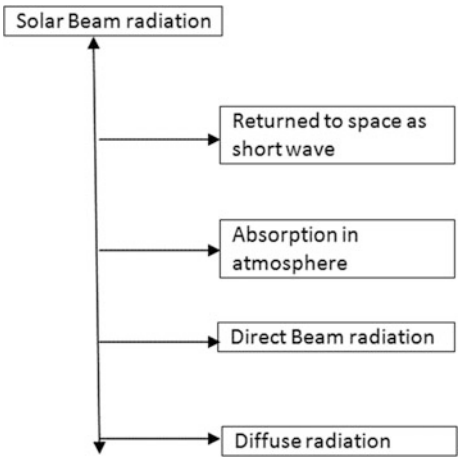


Fig. 1 Effects occurring as ethereal solar radiation is incident upon the atmosphere

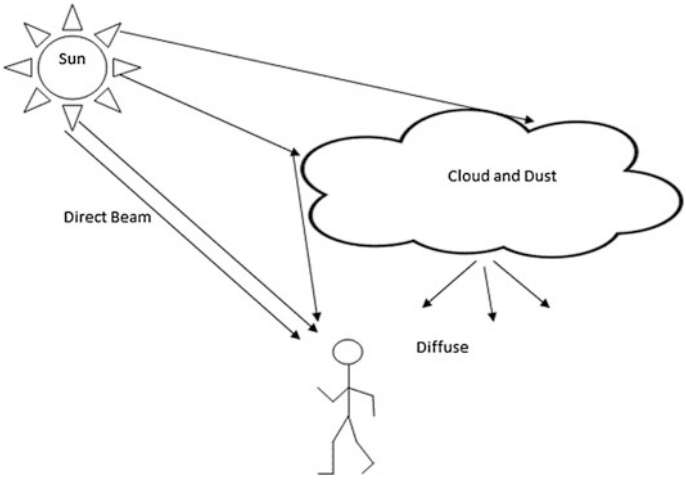


Fig. 2 Pictorial representation of direct beam and diffuse radiation

2 Photovoltaic Model and Its Governing Equation

The popularly known characteristics of a standard silicon p–n junction are shown in Fig. 3.

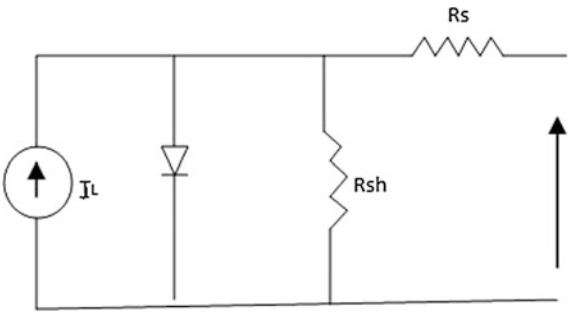
Mathematically, this can be given by

$$I = I_0\{\exp(V/V_T) - 1\}, \tag{1}$$

where I is the solar cell current, I_0 is the current in reverse saturation, V is the solar cell voltage, V_T is known as voltage equivalent of temperature, and at room temperature of 20 °C its approximate value is 26 mV which is equal to KT/Q where K is the Boltzmann's constant, T is the temperature in °K, and Q is the charges of electron.

When the p–n junction is illuminated by radiant flux, the characteristics get improved in shape and shift downward as photon-induced component is added with reverse leakage current.

Fig. 3 Equivalent circuit of solar PV cell



Equation (1) is the diode equation which is modified as follows:

$$I = -I_{SC} + I_0\{\exp(V/V_T) - 1\}. \tag{2}$$

When the junction is short circuited at its terminal V becomes zero and a finite amount current $I = -I_{SC}$ flows through the external path, appearing from p-side. I_{SC} is short-circuit current and its magnitude depends on solar insulation.

Now, a voltage source is installed in external path of p-side, as the magnitude value of the voltage is continually increased from zero and starts decreasing the current. Mathematically, representation of the open-circuit voltage *Voc* can be given by

$$Voc = V_T \ln\{(I_{sc}/I_0) + 1\}. \tag{3}$$

Typically, $I_{SC} = 2\,\text{A}$, $I_0 = 1\,\text{mA}$, $V_{OC} = 0.55\,\text{A}$ (at room temperature). Mathematically, the I–V characteristic of solar cell may be written (as per standard sign convention) as

$$I = I_{SC} - I_0\{\exp(V/V_T) - 1\}. \tag{4}$$

Solar energy is a best alternative for electric power generation and electromagnetic solar radiation which is exactly converted into electrical energy by solar PV module. Basically, solar module is made up of string of solar cell [8] and solar cell is an ordinary p–n junction diode of semiconductor. The PV module is composed of silicon cells. In open-circuit condition, the silicon cells have operating voltage of 0.7 V. The current ratings depend on the area of the particular cell. The current ratings of the module are proportional to the area of the solar cell, to obtain the higher power output. The solar photovoltaic modules are connected in series and parallel combinations to form solar PV array.

3 Equivalent Circuit

The I–V characteristic given in Eq. (4) is determined for ideal condition, considering internal series resistance of cell is zero and the shunt resistance is infinite. In actual practice, both have finite values, which would alter the characteristics. The ideal and practical equivalent circuits of the solar cell are shown in Fig. 3.

In practice I_{SC} is no longer equal to light generated by current I_V but is less by shunt current through shunt resistance, R_{sh}. The P–V current equation of solar P–V cell can be modified as follows:

$$I = I_L - I_0\{\exp(V + IR_S)/V_T - 1\} - (V + IR_S)/R_{sh} \tag{5}$$

V is P–V voltage, I_L is the photoelectric current (Figs. 4, 5 and 6).

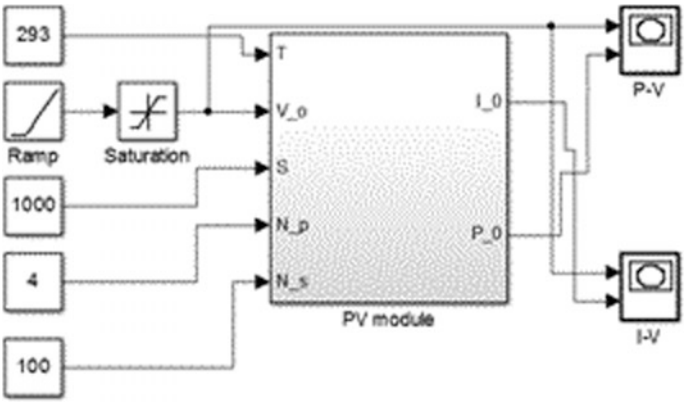


Fig. 4 Simulation model of PV cell

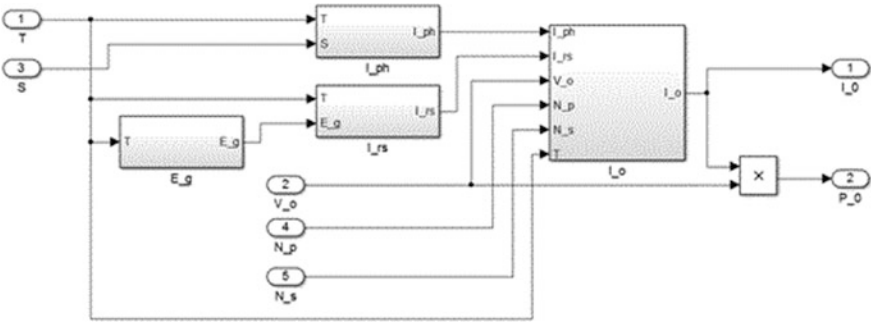


Fig. 5 Operational functional sub-block of PV model

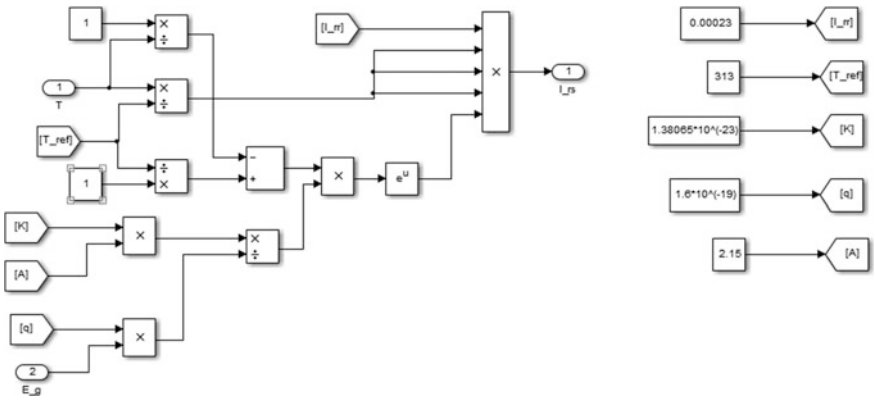


Fig. 6 Operational functional sub-block of PV model

For typical high-quality, one-square-inch silicon cell, series resistance, R_s = 0.05–0.10 Ω, shunt resistance, R_{sh} = 200–300 Ω.

T temperature of the module in Kelvin
I P–V current
I_L I_{SC}.

4 Simulation Model of PV Array with Its P–V and V–I Characteristics

The simulated PV model with four PV modules connected in series and its output of P–V and V–I characteristics are shown in Fig. 1. The programming-based code simplifies the simulation block and increases its efficiency and readability. The PV simulation model has been implemented through different stages. In Simulink model all the stages (functional blocks and sub-blocks) are interrelated. Simulation parameters are listed in Table 1. Simulation program of PV module has been implemented by series–parallel combination using some simulation parameter constant temperature and varying solar irradiance [7] and its output of V–I and P–V characteristics as shown in Figs. 7 and 8.

The plot of electrical characteristics of a PV module is represented at a particular temperature and solar irradiance. This is plotted on a graph with voltage on the axis of abscissas and current on the axis of ordinates, while maintaining irradiance and temperature levels constant. I–V curve presents an unlimited number of current–voltage operating points; the distinct operating point has been determined by the electric load associated to a PV system.

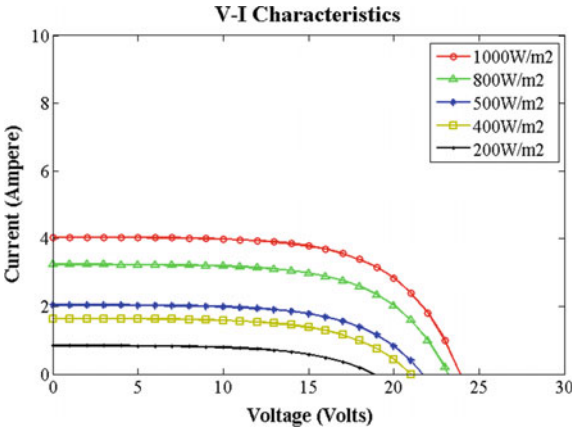


Fig. 7 V–I characteristic of PV cell

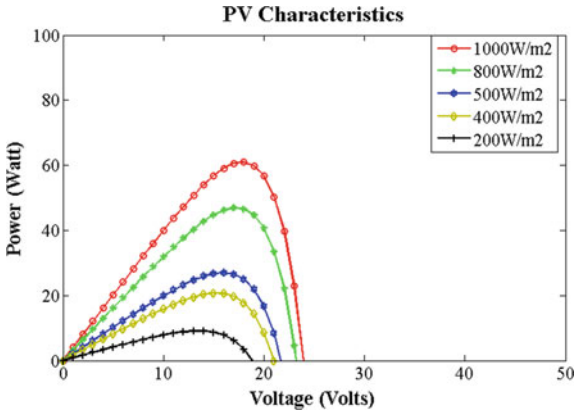
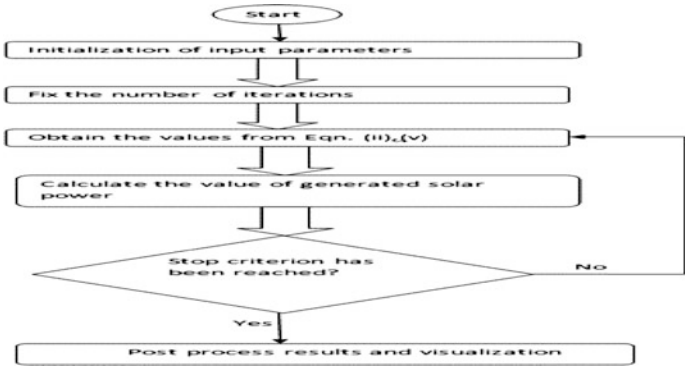


Fig. 8 V–P characteristic of PV cell

5 Matlab Coding Output of PV Array with Its P–V and V–I Characteristics



See Figs. 7, 8 and Table 1.

6 Conclusion

In summary, this paper dealt with PV module characteristics under two different approaches. An analytical model presented under consider of Matlab Simulink method with different solar irradiations and varying temperatures. Also this analytical method should implement through Matlab coding with different irradiances and temperatures. This result is obtained in the form of two fundamental graphs V–I and P–V characteristics.

Acknowledgements We would like to thank Jadavpur University, Kolkata, India for providing all the requisite help to carry out this work. This work gives assistance to the program of DRS, UPE II, UGC, Govt. of India (GOI).

References

1. Ma, T., Yang, H.X., Lu, L.: Development of a model to simulate the performance characteristics of crystalline silicon photovoltaic modules/strings/arrays. Sol. Energy **100**, 31–41 (2014)
2. Patel, H., Agarwal, V.: MATLAB-based modeling to study the effects of partial shading on PV array characteristics. IEEE Trans. Energy Convers. **23**, 302–310 (2008)
3. Deshkar, S.N., Dhale, S.B., Mukherjee, J.S., Babu, T.S., Rajasekar, N.: Solar PV array reconfiguration under partial shading conditions for maximum power extraction using genetic algorithm. Renew. Sustain. Energy Rev. **43**, 102–110 (2015)
4. Fezzani, A., Mahammed, I.H., Said, S.: MATLAB-based modeling of shading effects in photovoltaic arrays. In: 15th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering—STA2014, Hammamet, Tunisia, pp. 781–787 (2014)
5. Rahman, M.Z.: Status of Selective emitters for the p-type c-Si solar cells. Sci. Res. Int. J. Opt. Photon. **2**, 129–134 (2012)
6. Bai, J., Cao, Y., Hao, Y., Zhang, Z., Liu, S., Cao, F.: Characteristic output of PV systems under partial shading or mismatch conditions. Sol. Energy **112**, 41–54 (2015)
7. Shukla1, A., Khare, M., Shukla, K.N.: Modeling and simulation of solar PV moduleon MATLAB/simulink, pp 18516–18527 (2015)
8. Ramaprabha, R., Mathur, B.L.: Impact of partial shading on solar PV module containing series connected cells. Int. J. Recent Trends Eng. **2**(7), 56–60 (2009)
9. Ji, Y.H., Kim, J.G., Park, S.H., Kim, J.H., Won, C.Y.: C-language based PV array simulation technique considering effects of partial shading. In: IEEE International Conference on Industrial Technology, ICIT, pp. 1–6 (2009)
10. Kasera, J., Chaplot, A., Maherchandani, J.K.: Modeling and simulation of wind-PV hybrid power system using MATLAB/simulink. In: IEEE Students' Conference on Electrical, Electronics and Computer Science, pp. 1–4 (2012)

Determination of Yield Coefficients of Methane and Carbon Dioxide in Methanogenesis to Predict the Overall Performance of a Biogas Digester

Joyoti Biswas, Ranjana Chowdhury and Pinaki Bhattacharya

Abstract Biogas has globally remained a renewable energy source derived from anaerobic digestion of organic matter. This process treats the organic matter efficiently producing a effluent whose residue is rich in inorganic elements like nitrogen and phosphorous needed for effective plant growth. In our investigation determination of yield coefficients of methane and carbon dioxide in methanogenesis was done considering volatile fatty acid and long chain fatty acids as the principal substrates obtained after the acidogenic degradation of vegetable market waste. The purpose of considering these two substrates is that these are the primary substrates for the methanogens needed to produce biogas. Besides, a pure chemical source of volatile fatty acid and long chain fatty acids has been used to conveniently determine yield coefficients of the products of methanogenesis which matched well with the biogas generated from the digester.

Keywords Biogas · Batch mode · Biodigester · Yield coefficients

1 Introduction

Cattle excreta has been used in many rural areas of underdeveloped and developed countries for generation of biogas [1]. Use of solid market waste presents an alternative potential source for generation of biogas. The rural population which entirely depended on firewood could now depend on an alternative source of energy, namely, biogas. It is an inexpensive fuel that can be used to replace the various fossil fuels like coal petroleum, etc. [2].

J. Biswas (✉)
Department of Chemistry, Dr. Sudhir Chandra Sur Degree Engineering College, Kolkata, India
e-mail: jaysagnik@yahoo.com

R. Chowdhury · P. Bhattacharya
Department of Chemical Engineering, Jadavpur University, Kolkata, India

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_3

Biogas, a product of anaerobic digestion, essentially is a mixture of methane and carbon dioxide with a trace amount of some other gases like hydrogen sulfide. Methane formation in anaerobic digestion involves a complex array of microorganisms broadly divided into acidogenic and methanogenic bacterial consortia. The efficiency of biogas generation will depend on proper coordination of acidogenic and methanogenic consortia [3].

Most of the works till date have been done regarding the overall performance of the digester taking into consideration various parameters, like, pH, temperature, mixing, substrate, and hydraulic retention time [4]. Work related to isolation of bacterial consortia of the two phases, namely acidogenesis and methanogenesis, has not yet been reported. Taking the above-mentioned fact into consideration, growth kinetics of the isolated bacterial consortia along with the yield coefficients of the two important constituents of biogas have been studied in laboratory scale to predict the overall biogas generation rate in bio-digesters. Growth kinetics of acidogenic and methanogenic strains already isolated has been determined and reported earlier [5].

In the present study yield coefficients of the major products of biogas, namely, methane and carbon dioxide, have been studied in batch mode using Erlenmeyer flasks under controlled conditions using volatile fatty acid (VFA) and long chain fatty acid (LCFA) as substrate.

2 Materials and Methods

2.1 Substrate

Municipal market waste along with the chemical sources of volatile fatty acids and long chain fatty acids has been used as substrate in our present study. The waste which consisted of harmful pathogens, after collection, was subjected to sterilization at 121 °C for 15 min followed by electric grinding. Sterilization as well as electrical grinding reduces the volume of the waste which was then mixed with distilled water to obtain a desired consistency and concentration level. The substrate so formed was hand sorted to remove the plastics and metals from it.

2.2 Separation of Methanogenic Bacterial Consortia from the Reaction Broth

Inoculum for methanogenesis was collected from a running biogas digester when the production of biogas occurred at a substantial rate. Further to get the bacterial strains adapted to the present condition, the inoculum was enriched in a medium

containing 0.82 kg/m^3 of pure chemical sources of VFA, namely, sodium salts of acetic, propionic butyric and valeric acids in equal proportion, and a pure source of LCFA, namely, sodium oleate at a concentration of 0.08 kg/m^3. To obtain a better yield, monitoring of pH was done and maintained at 6.8–7.0. Maintenance of the bacterial culture at a desired activity level was done by repeated culture method in general bacterial medium.

2.3 Preparation of the Seed Culture for the Digester

Inoculum for operation of the digester was prepared using the reaction broth of a running cow dung-based biogas plant. A measured volume of 20 ml general sterile bacterial medium was mixed with 1 g of the inoculum collected from the digester and was allowed to grow for 2 days in an incubator at 40 °C. To replicate the actual condition of the digester, inoculation of selective medium having the composition of feed slurry of 20 kg/m^3 concentration was done with the suspended bacterial culture grown in general bacterial medium. The activated state of the bacterial consortia was obtained by enrichment in the selective medium for at least 5 times.

2.4 Analysis of the Biogas

Composition of the major products of biogas was analyzed by gas chromatograph (CHEMITO, MODEL No. 8510) using thermal conductivity detector (TCD) with a PorapaK Q column. The carrier gas flow rate of hydrogen was maintained at 20 ml/min. The detector, injector, and oven temperatures were maintained at 80 °C, 80 °C, and 60 °C, respectively [6].

3 Theoretical Analysis

The reaction scheme for methanogenesis to predict the yield of methane and carbon dioxide in biogas is as follows [7].

The scheme is as follows:

$$\text{VFA} + \text{Biomass} \rightarrow \text{CH}_4 + \text{CO}_2 + \text{Biomass} \quad (1)$$

$$\text{LCFA} + \text{Biomass} + \text{CO}_2 \rightarrow \text{VFA} + \text{CH}_4 + \text{Biomass}. \quad (2)$$

The yield coefficients of the products of methane and carbon dioxide in biogas can be calculated using the relation as follows:

$$\mathbf{Y_{p/x} = \frac{mass\ of\ products\ formed}{increase\ in\ cell\ mass}}.$$

4 Results and Discussion

4.1 Methanogenesis

Biomass concentration during methanogenesis has been obtained as a function of time with variable initial VFA and LCFA concentrations as a parameter. The cell growth time history during methanogenesis has been formulated using the experimental results which have been depicted in Figs. 1 and 4. The concentrations of the major products in methanogenesis, namely, methane and carbon dioxide, obtained as a function of time with varying initial concentration of VFA and LCFA as parameter have been plotted in Figs. 2, 3, 5, and 6.

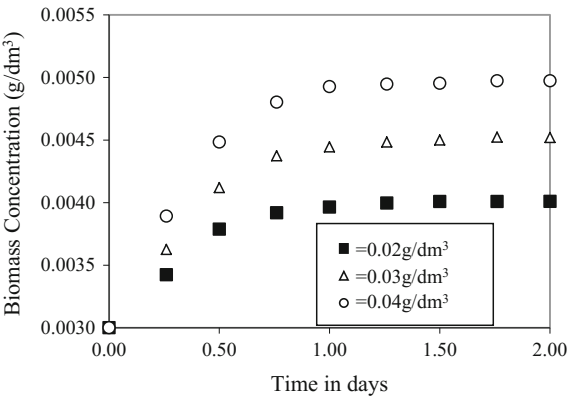


Fig. 1 Experimental concentration time histories of biomass during methanogenic degradation of acidogenically degraded (2 days) vegetable slurry in Erlenmeyer flasks using VFA concentration as a parameter

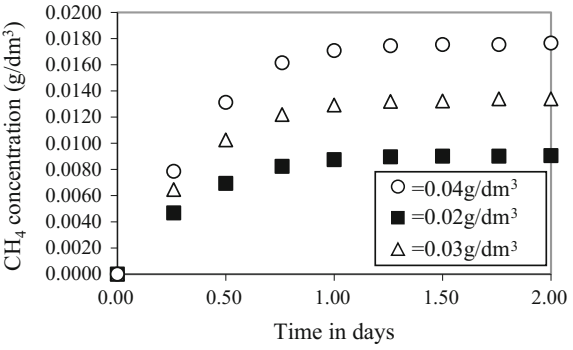


Fig. 2 Experimental concentration time histories of CH_4 during methanogenic degradation of acidogenically degraded (2 days) vegetable slurry in Erlenmeyer flasks using VFA concentration as a parameter

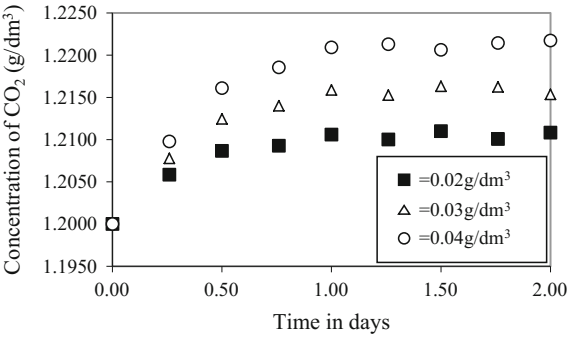


Fig. 3 Experimental concentration time histories of CO_2 during methanogenic degradation of acidogenically degraded (2 days) vegetable slurry in Erlenmeyer flasks using VFA concentration as a parameter

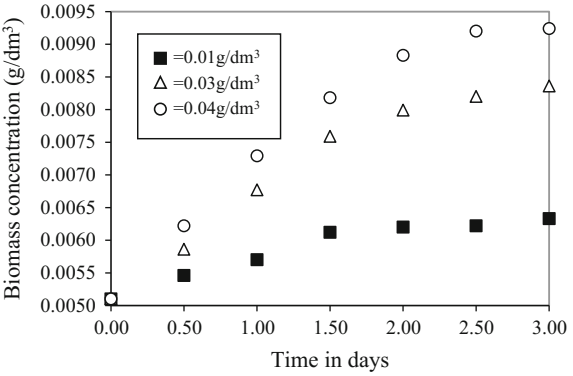


Fig. 4 Experimental concentration time histories of biomass during methanogenic degradation of acidogenically degraded (2 days) vegetable slurry in Erlenmeyer flasks using LCFA concentration as a parameter

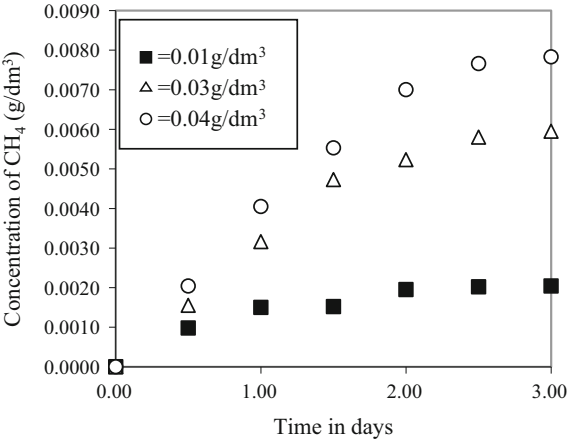


Fig. 5 Experimental concentration time histories of methane during methanogenic degradation of acidogenically degraded (2 days) vegetable slurry in Erlenmeyer flasks using LCFA concentration as a parameter

Figures 1 and 4 clearly depict the life cycle of the methanogenic strains comprising exponential and stationary phases. The lag phase could not be seen in the figure as the constant enrichment process has been able to maintain the bacterial

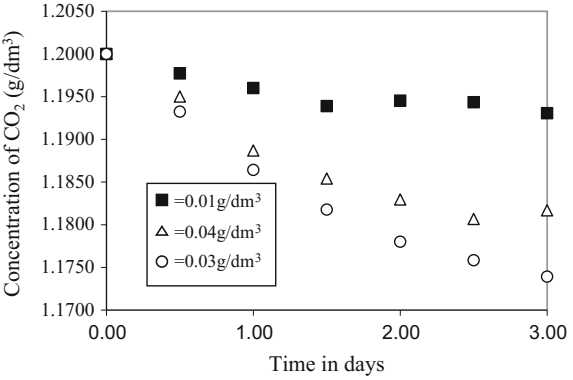


Fig. 6 Experimental concentration time histories of CO_2 during methanogenic degradation of acidogenically degraded (2 days) vegetable slurry in Erlenmeyer flasks using LCFA concentration as a parameter

Table 1 Yield coefficients obtained using batch mode of operation

Reaction type	Yield coefficients ($Y_{p/x}$)	
	CO_2	CH_4
VFA methanogenesis	11.128	9.522
LCFA methanogenesis	6.289	1.905

strains at the desired activity level and as a consequence the adaptability time needed for acclimatization to the new environment was not needed by the microorganisms. The yield coefficients of CH_4 and CO_2 obtained from VFA and LCFA methanogenesis are presented in Table 1.

5 Conclusion

An elaborate and detailed study was carried out regarding the improvement of the yield of biogas generation rate from municipal market waste. Initially, laboratory scale study in Erlenmeyer flask was carried out to ensure the feasibility of producing biogas from municipal market waste.

After establishing the feasibility of biogas production from the municipal market waste, the laboratory scale study was further carried out to determine the yield coefficients of methane and carbon dioxide using batch mode under controlled conditions. By proper optimization of the operating conditions, the methanogenic strains were isolated and enriched in the respective media to ensure proper activation level of the microorganisms. While conducting this study pure sources of VFA and LCFA have been used as a limiting substrate. Eventually, the yield coefficients obtained for methane and carbon dioxide in VFA and LCFA methanogenesis predicted successfully the yield of methane and carbon dioxide in 10 dm^3 biogas digester [8]. The percentage composition of methane in biogas obtained from the digester was as high as 70–75%.

Acknowledgements This study was done under UGC major project scheme and the authors acknowledge the financial support granted under the scheme.

References

1. Home, T.: Case study: the team process for biomethanation of organic solid wastes. 1–7 (2002)
2. Khanolkar, V.D., Dalvi, S.D.: Treatment of bagasse black liquor at pudumjee pulp and paper mills ltd., Pune. Bio Energy News **4**, 15–19 (2000)
3. Tekin, A.R., Dalgic, A.C.: Biogas production from olive pomace. Resour. Conserv. Recycl. **30**, 301–313 (2000)
4. Callaghan, D.A., Thayanithy, J., Forster, C.F.: Continuous Co digestion of cattle slurry with fruit and vegetable wastes and chicken manure. Biomass Bioenergy **27**, 71–77 (2002)
5. Biswas, J., Chowdhury, R., Bhattacharya, P.: Determination of growth kinetics of isolated consortia for biogas generation. In: IEEE (ICONCE 2014), pp. 123–125
6. Angelidaki, I., Peterson, S.P., Ahring, B.K.: Effects of lipids on thermophilic anaerobic digestion and reduction of lipid inhibition upon addition of bentonite. Appl. Microbiol. Biotechnol. **33**, 469–472 (1990)
7. Angelidaki, I., Ellegard, L., Ahring, B.K.: A comprehensive model of anaerobic bioconversion of complex substrates to biogas. Biotechnol. Bioeng. **63**, 560–564 (1999)
8. Biswas, J., Chowdhury, R., Bhattacharya, P.: Kinetic studies of biogas generation using municipal waste as feed stock. Enzyme Microb. Technol. **38**, 493–503 (2006)

A Novel Algorithm for Economic Load Dispatch Using a New Optimization Technique

S. Mandal, G. Das, M. De, B. Tudu and K.K. Mandal

Abstract Economic load dispatch is one of fundamental issues in optimal power system operation. Several classical and modern heuristic techniques have been used to solve the problem. A novel algorithm is presented in this paper for solving economic load dispatch problem using a recently introduced simple yet powerful optimization technique called Jaya algorithm (JA). Most of the modern heuristic techniques require setting of control parameters which are usually problem specific. Moreover, there is no specific rule for the selection of these user-defined control parameters for most of the heuristic techniques. A wrong parameter may even lead to premature convergence. One of the major advantages in the proposed algorithm is that it does not require any problem-dependent control parameters. The economic load dispatch (ELD) problem is formulated by considering prohibited zones, ramp rate limits, and transmission losses satisfying a set of equality and inequality constraints. The proposed algorithm is tested on a six-unit test system in order to verify its effectiveness and efficiency. The simulation results are compared with those obtained by modern heuristic techniques. The simulation results show that the proposed technique has the capability of producing comparable results.

Keywords Economic load dispatch · Ramp rate limits · Prohibited operating · Zone · Loss · Jaya algorithm

1 Introduction

Economic load dispatch (ELD) is one of the crucial problems in power system functioning. The aim of economic load dispatch (ELD) is to schedule the committed generators in such a way that total fuel cost is minimum fulfilling the required load

S. Mandal
Department of Electrical Engineering, Jadavpur University, Kolkata 700032, India

G. Das · M. De · B. Tudu · K.K. Mandal (✉)
Department of Power Engineering, Jadavpur University, Kolkata 700098, India
e-mail: kkm567@yahoo.co.in

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_4

demand under a set equality and inequality constraints. As the generation cost for thermal power is excessively high due high input cost, an optimum dispatch program for the committed generators can save considerable amount of revenue. The problem becomes more complex if one is required to consider valve point effects, transmission losses, ramp rate limits (RRL), prohibited operating zones (POZ), and other constraints. Thus, for a practical system, it becomes a complex nonlinear optimization problem formulation. A wide variety of methods including modern heuristic optimization techniques have been applied to solve this problem of economic load dispatch. Different mathematical and classical optimization techniques have been used to formulate the ELD problem. These comprise dynamic programming (DP), quadratic programming (QP), linear programming, homogenous linear programming, nonlinear programming techniques, and interior point method approaches [1–3]. In order to make numerical methods more efficient, artificial intelligent techniques, like the Hopfield neural networks, have been successfully employed to solve ELD problems [4]. Several modern heuristic techniques have been successfully applied in recent years to solve economic load dispatch problems. Some of them include genetic algorithm [5], evolutionary algorithms [6], artificial neural network [7], simulated annealing [8], ant colony optimization [9], and differential evolution [10]. Park et al. [11] presented an improved particle swarm technique for solving economic dispatch problem with non-smooth cost function. Xiong et al. [12] proposed a novel algorithm using biogeography-based optimization technique for economic dispatch problem. A chaotic differential evolution-based method was proposed by Lu et al. [13] for solving dynamic ELD problem and represented encouraging results. Recently, Cuckoo search algorithm was successfully applied to solve economic dispatch problem by several researchers [14]. A hybrid optimization technique using differential evolution and particle swarm optimization technique was proposed by Parouha and Das [15]. The proposed algorithm was tested on several systems and encouraging results were presented. More recently, an algorithm using colonial competitive differential evolution technique was proposed by Ghasemi et al. [16] for solving optimal economic load dispatch problem considering several constraints.

Jaya algorithm (JA) is one of the most recent population-based optimization techniques and was proposed by Rao in 2016 [17]. It is a simple, easy to implement, yet very powerful optimization algorithm for both constrained and unconstrained optimization problems. It has been tested on several constrained and unconstrained engineering optimization problems successfully [18]. This paper proposes a novel optimization technique using Jaya algorithm for solving economic dispatch problem with ramp rate limits (RRL), prohibited operating zones (POZ), and losses satisfying a set of other equality and inequality constraints. The simulation results are presented and compared with other modern techniques. It is found that the proposed technique is capable in producing comparable results.

2 Problem Formulation

In this section, the problem for economic load dispatch (ELD) is described. The primary goal of ELD problem is to reduce the total fuel cost of generating units meeting the system requirement under some operating constraints. The fuel cost characteristic for any unit is concluded to be approximated by segments of quadratic functions of the active power output of the generator for clarity. Thus, the problem may be narrated as the minimization of the total fuel cost and defined by (1)

$$\text{Minimize} \quad FC(P_g) = \sum_{i=1}^{N_g} (a_i P_{gi}^2 + b_i P_{gi} + c_i), \tag{1}$$

where $FC(P_g)$ represents the total fuel cost of generation ($/h), a_i, b_i, c_i are the coefficients of fuel cost of the ith generating unit, P_{gi} is the power generated by the ith unit, and N_g represents the number of thermal units.

The fuel cost is to be reduced subject to the following equality and inequality constraints.

(i) Upper and lower limit of power generation

From practical point of view, power generated by every unit must be bounded by specified maximum and minimum limits. This can be specified as follows:

$$P_{gi}^{\min} \leq P_{gi} \leq P_{gi}^{\max}, \tag{2}$$

where $P_{gi}^{\min}$ is the minimum power generated by ith unit and $P_{gi}^{\max}$ is the maximum power generated by the ith unit.

(ii) Power balance constraints

Sum of the power produced by all the units must be equal to sum of power demand and losses. This can is expressed as follows:

$$\sum_{i=1}^{n} P_{gi} = P_D + P_L, \tag{3}$$

where P_D is the total power demand and P_L is the total transmission loss.

The transmission loss P_L can be determined using B matrix and is given by (4):

$$P_L = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P_i B_{ij} P_j, \tag{4}$$

where B_{ij}'s are the elements of loss coefficient matrix.

(iii) Ramp Rate Limits (RRL)

To adjust power generated by each unit so as total demand is met, power generation by the units is to be increased or decreased gradually and continuously and this is limited by the following:

(a) as generation increases

$$P_{gi} - P_{gi}^{0} = \leq UR_i. \tag{5}$$

(b) as generation decreases

$$P_{gi} - P_{gi}^{0} = \leq DR_i, \tag{6}$$

where P_{gi} and P_{gi}^{0} are the current power generation and previous power generation, respectively. UR_i is the up ramp limit of the ith generator (MW/time period) and DR_i is the down ramp limit of the ith generator (MW/time period).

(iv) Prohibited operating zones (POZ)

Some of the generating units may have certain regions in their cost curve where generation is either not desirable or not possible due physical limitations. This can be expressed as

$$\begin{aligned} &P_{gi}^{\min} \leq P_{gi} \leq P_{gi,1}^{l} \\ &P_{gi,j-1}^{u} \leq P_{gi} \leq P_{gi,j}^{l} \quad j = 2, 3, \ldots, n_i, \\ &P_{gi,n_i}^{u} \leq P_{gi} \leq P_{gi}^{\max} \end{aligned} \tag{7}$$

where n_i is the number of prohibited zones of ith unit, $P_{gi,j}^{l}$ is the lower active power limit of prohibited zone j of the ith unit (MW), and $P_{gi,j-1}^{u}$ is the upper active power limit of prohibited zone $j - 1$ of the ith unit (MW).

3 Jaya Algorithm

Recently, a new powerful optimization technique was introduced by Rao in 2016 [18]. It is a population-based heuristic optimization technique called Jaya algorithm (JA). The main advantages of the algorithm are simple, easy to implement and does not require user-defined control parameters which are usually problem specific. Like other population-based techniques, it starts with the randomly generated initial candidate solutions. Let $f(x)$ be the objective function to be minimized (for the

present case) and the best candidate solution and worst candidate solution for the entire candidate solutions be represented by $f(x)_{best}$ and $f(x)_{worst}$, respectively. If $X_{j,k,i}$ is the jth variable for the kth candidate during ith iteration, then it is modified as follows [19]:

$$X^{M}_{j,k,i} = X_{j,k,i} + r_{1,j,i}\left(X_{j,best,i} - \left|X_{j,k,i}\right|\right) - r_{2,j,i}\left(X_{j,worst,i} - X_{j,k,i}\right), \tag{8}$$

where $X_{j,best,i}$ and $X_{j,worst,i}$ represent the value of the jth variable for the best candidate and worst candidate, respectively. $r_{1,j,i}$ and $r_{2,j,i}$ are the two random numbers in the range (0, 1) for the jth variable during the ith iteration. $X^{M}_{j,k,i}$ is the modified updated value of $X_{j,k,i}$. The second term on the right-hand side of (8) indicates the tendency of the candidate solution to move closer to the best value, while the last term indicates the tendency to avoid the worst solution. The fitness or objective function is evaluated with both the original (i.e., $X_{j,k,i}$) and modified value (i.e., $X^{M}_{j,k,i}$). The terms which give the better value (minimum for the present case) of the objective function are selected and stored. These terms become the input for the next solution. Two random numbers r_1 and r_2 in the range of (0, 1) help to search the entire solution and improve the exploration capability of the algorithm. The proposed algorithm can be described and implemented as follows:

Step 1 Initialize population, maximum iteration number; generate initial solution randomly satisfying the problem constraints.
Step 2 Calculate the fitness function according to (1). Find the best and worst solutions in the entire population.
Step 3 Modify solutions using (8) satisfying the constraints.
Step 4 If the solution corresponding to $X^{M}_{j,k,i}$ is better than that corresponding to $X_{j,k,i}$, replace the previous solution by the present solution. Otherwise retain the previous solution.
Step 5 If the termination criterion, i.e., the maximum iteration, is reached for the present study then go to next step else go to Step 2.
Step 6 Report the optimal solution.

4 Results and Discussion

The proposed algorithm is applied on a six-generator test system [19] to verify its effectiveness and was implemented using Matlab on 3.0 GHz, 4.0 GB RAM PC. The power generated by each unit is chosen as the control variables for the present problem. Table 1 shows the generator capacity and fuel cost coefficients. Ramp rate limits and prohibited zones for the generating units are shown in Table 2. Valve point effect is neglected for a direct comparison for the present system. Other data including B-loss coefficient matrix is taken from [19]. The population size is taken as 40 after several runs.

Table 1 Generating unit's capacity and cost coefficients

Unit	$P_{gi}^{\min}$	$P_{gi}^{\max}$	a_i	b_i	c_i
1	100	500	0.0070	7.00	240
2	50	200	0.0095	10.00	200
3	80	300	0.0090	8.50	220
4	50	150	0.0090	11.00	200
5	50	200	0.0080	10.50	220
6	50	120	0.0075	12.00	190

Table 2 Ramp rate limits and prohibited zones of generating units

Unit	P_{gi}^{0}	UR_i (MW/h)	DR_i (MW/h)	Prohibited zones (MW)
1	440	80	120	[210 240] [350 380]
2	170	50	90	[90 110] [140 160]
3	200	65	100	[150 170] [210 240]
4	150	50	90	[80 90] [110 120]
5	190	50	90	[90 110] [140 150]
6	110	50	90	[75 85] [100 105]

Table 3 Simulation results with optimal generation loss and cost

Unit (MW)	Demand (MW)		
	800	1100	1263
P1 (MW)	328.0082	414.1539	451.4248
P2 (MW)	85.4070	139.7254	176.0929
P3 (MW)	171.9176	241.6763	255.8996
P4 (MW)	60.0000	107.9083	150.0000
P5 (MW)	110.0044	137.7718	174.2446
P6 (MW)	50.0000	68.4002	67.7409
Total generation (MW)	8.3892	1109.6359	1275.4028
Losses (MW)	5.3892	9.6359	12.4028
CPU time (s)	13.21	13.15	13.18
Iterations	200	200	200
Fuel cost ($/h)	9,433.45	13,278.51	15,448.74

The results including optimal generation, optimal cost, and loss are shown in Table 3 for the demand of 800, 1100, and 1263 MW. For each of the demand 50 trial runs were made and it was noticed that almost 100% success is achieved. Maximum iteration was set at 200 after several trial runs. It was increased in step of 50 and no improvement in results was noticed beyond 200.

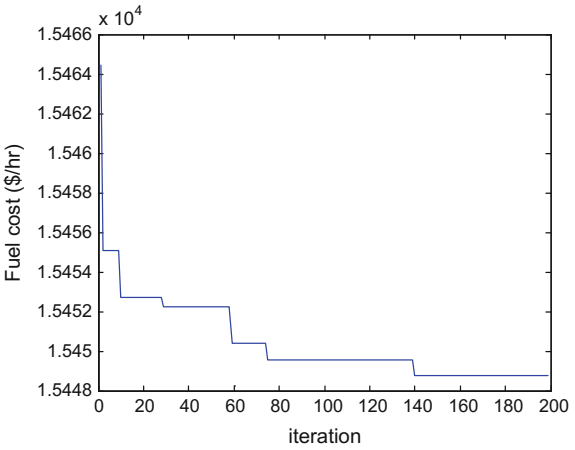


Fig. 1 Convergence characteristics for minimum fuel cost

Table 4 Comparison of results with different methods

	Proposed method	PSO [19]	Genetic algorithm [19]
Fuel cost ($)	15,448.74	15,450.00	15,459.00
Loss (MW)	12.4028	12.9584	13.0287

It is observed from Table 3 that optimal fuel cost is $15448.74 for a demand of 1263 MW while the loss is found to be 12.4048 MW. The computation tine is 13.18 s. The voltages for all the generating units are well within the limits. The convergence characteristic for the optimal fuel cost for a load demand of 1263 MW is shown in Fig. 1.

The results are also compared with other methods like particle swarm optimization technique (PSO) and genetic algorithm (GA) [19] and the comparison result is shown in Table 4 for the demand of 1263 MW. It is clearly observed that the proposed method is capable of producing improved results.

5 Conclusion

Economic load dispatch (ELD) is most significant issues in present-day power system operation. The basic aim is to minimize fuel cost while fulfilling the various constraints. The economic load dispatch is a complex optimization problem and many optimization techniques and algorithms have been used to solve the problem. A new algorithm is proposed in the present paper using a recently proposed new optimization technique. The proposed algorithm is applied on a six-generator test system in order to verify its effectiveness. The simulation results show that it has the capability to produce good quality solutions. The results obtained by the proposed algorithm are also compared with other methods, like particle swarm optimization

technique (PSO) and genetic algorithm (GA), and comparison results are presented. It is found the proposed algorithm can produce superior results in respect of fuel cost and loss.

References

1. Yang, H.T., Chen, S.L.: Incorporating a multi-criteria decision procedure into the combined dynamic programming/production simulation algorithm for generation expansion planning. IEEE Trans. Power Syst. **4**(1), 165–175 (1989)
2. Liang, Z.X., Glover, J.D.: A zoom feature for a programming solution to economic dispatch including transmission losses. IEEE Trans. Power Syst. **7**(3), 544–550 (1992)
3. Jabr, R.A., Coonick, A.H., Cory, B.J.: A homogeneous linear programming algorithm for the security constrained economic dispatch problem. IEEE Trans. Power Syst. **15**(3), 930–936 (2000)
4. Su, C.T., Chiou, G.J.: A fast computation Hopfield method to economic dispatch of power systems. IEEE Trans. Power Syst. **12**(4), 1759–1764 (1997)
5. Chen, P.H., Chang, H.C.: Large-scale economic dispatch by genetic algorithm. IEEE Trans. Power Syst. **10**, 1919–1926 (1995)
6. Sahoo, S., Mahesh Dash, K., Prusty, R.C., Barisal, A.K.: Comparative analysis of optimal load dispatch through evolutionary algorithms. Ain Shams Eng. J. **6**, 107–120 (2015)
7. Kumar, J., Seblé, G.B.: Clamped state solution of artificial neural network for real-time economic dispatch. IEEE Trans. Power Syst. **10**(2), 925–931 (1995)
8. Wong, K., Fong, C.: Simulated annealing based economic dispatch algorithm. IEEE Proc. Gener. Transm. Distrib. **140**(6), 509–515 (1993)
9. Song, Y.H., Chou, C.S., Stonham, T.J.: Combined heat and power economic dispatch by improved ant colony search algorithm. Electr. Power Syst. Res. **52**(2), 115–121 (1999)
10. Coelho, L.D.S., Mariani, V.C.: Improved differential evolution algorithms for handling economic dispatch optimization with generator constraints. Energy Convers. Manage. **48**, 1631–1639 (2007)
11. Park, J.B., Jeong, Y.W., Shin, J.R., Lee, K.Y.: An improved particle swarm optimization for nonconvex economic dispatch problems. IEEE Trans. Power Syst. **25**(1), 156–166 (2010)
12. Xiong, G., Shi, D., Duan, X.: Multi-strategy ensemble biogeography-based optimization for economic dispatch problems. Appl. Energy **111**, 801–811 (2013)
13. Lu, Y., Zhou, J., Qin, H., Wang, Y., Zhang, Y.: Chaotic differential evolution methods for dynamic economic dispatch with valve-point effects. Eng. Appl. Artif. Intell. **24**, 378–387 (2011)
14. Ponnusamy, A., Rengarajan, N.: Optimal power flow solution using cuckoo search algorithm. ARPN J. Eng. Appl. Sci. **9**(12), 2687–2691 (2014)
15. Parouha, R.P., Das, K.N.: A novel hybrid optimizer for solving Economic Load Dispatch problem. Electr. Power Energy Syst. **78**, 108–126 (2016)
16. Ghasemi, M., Taghizadeh, M., Ghavidela, S., Abbasian, A.: Colonial competitive differential evolution: An experimental study for optimal economic load dispatch. Appl. Soft Comput. **40**, 342–363 (2016)
17. Rao, R.V.: Jaya: a simple and new optimization algorithm for solving constrained and unconstrained optimization problems. Int. J. Ind. Eng. Comput. **7**, 17–34 (2016)
18. Rao, R.V., More, K.C., Taler, J., Ocłon, P.: Dimensional optimization of a micro-channel heat sink using Jaya algorithm. Appl. Therm. Eng. **103**, 572–582 (2016)
19. Gaing, Z.L.: Particle swarm optimization to solving the economic dispatch considering the generator constraints. IEEE Trans. Power Syst. **18**(3), 1187–1195 (2003)

Investigating Economic Emission Dispatch Problem Using Improved Particle Swarm Optimization Technique

Meenakshi De, Gourab Das, S. Mandal and K.K. Mandal

Abstract This paper presents utilization of particle swarm optimization in solving combined economic emission dispatch (EED) problem. The economic emission dispatch is an important problem in power sector as it combines two major objectives viz., cost minimization and emission minimization while maintaining operational constraints. Several meta-heuristic techniques have been developed in recent times and have been applied on power dispatch problems. PSO is such a meta-heuristic technique where time-varying acceleration coefficients (TVAC) are incorporated and used in the EED problem in this work. Thus it addresses the techno-economic-environmental aspect of power system operation. Economic emission dispatch problem is first resolved using weighted sum method, and second trade-off curve between two objectives has been found, referred to as pareto front which traces solutions obtained by non-dominated approach of the problem. The formulation is implemented on IEEE 30 bus test system and outcome obtained validates effectiveness of this research work.

Keywords Economic emission dispatch · Time-varying acceleration coefficients incorporated particle swarm optimization (TVAC-PSO) · Meta-heuristic techniques · Non-dominated solutions

M. De · G. Das · K.K. Mandal (✉)
Department of Power Engineering, Jadavpur University, Kolkata, India
e-mail: kkm567@yahoo.co.in

M. De
e-mail: meen_230505@yahoo.in

S. Mandal
Department of Electrical Engineering, Jadavpur University, Kolkata, India

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_5

1 Introduction

Economic emission dispatch problem aims to satisfy dual objectives, i.e., cost minimization as well as emission minimization. This objective must be satisfied while maintaining all operational constraints within specified limits. Power dispatch problems mainly aim to allocate power to each of its generating units for finding the optimal solution. But increasing concerns for environment have focused on emission of harmful gases from the fossil-fired generators. Several methods are there for multi-objective optimization, such as analytical methods, meta-heuristic methods, etc. Recent advancements in meta-heuristics are note worthy, and application of such methods in this field shows promising results. A method of solving the EED problem is to incorporate emission as a constraint in a non-convex economic dispatch model [1]. Incorporation of renewable energy sources in the existing power system network is one approach for emission minimization, but it is not a cost-effective solution. Besides, including renewable energy sources requires extensive planning and therefore generally included in long-term goals [2]. Analytical approach is proposed in some literature for resolving the dispatch problem, but non-convexity in EED problem is not considered here [3]. Quadratic, linear programming approaches denote some conventional procedures for solving economic emission dispatch problem [4, 5], but they suffer from incapability of handling the nonlinear constraints [6]. The meta-heuristic methods utilized for solving the EED problem include genetic algorithm (GA) [7], particle swarm optimization (PSO) [8], differential evolution [9], ant optimization algorithm [10], etc. Improved PSO is utilized for this work with time-varying acceleration coefficients and used as an optimization tool.

2 Problem Formulation

Cost in generating units can be mathematically modeled accurately by considering the valve point loading effects.

$$F_{c,i}(P_i) = \sum_{i=1}^{N} a_i P_i^2 + b_i P_i + c_i + |e_i \times \sin(f_i(P_i^{\min} - P_i)) \tag{1}$$

$F_{c,i}$ is fuel cost function of N generators; a_i, b_i and c_i refers to cost coefficients of the ith generator, e_i and f_i represents valve point loading effect coefficients in ith generator. The valve point loading effects will increase number of local optima in formulation [11]. Emission level from generating units can be given by combining SO_x and NO_x emissions.

$$F_{e,i}(P_i) = \sum_{i=1}^{N} \alpha_i P_i^2 + \beta_i P_i + \gamma_i + \xi_i e^{\lambda i P i} \tag{2}$$

$F_{e,i}$ is the total emission level of N generating units; emission function coefficients of generating units are α_i, β_i, γ_i, ξ_i, λ_i. The power dispatch constraints are as following.

Equality constraints: This is given as follows:

$$\sum_{i=1}^{N} P_i - P_D - P_{loss} = 0 \tag{3}$$

Generation constraints: Each generator should produce power within lower operating limit and upper operating limit as follows:

$$P_i^{\min} \le P_i \le P_i^{\max},\ i \in \{1, 2, 3, \ldots, N\} \tag{4}$$

2.1 Weighted Sum Method

The weighted sum method attempts to transform set of objectives into single objective function [2]. Therefore, the combined economic/emission dispatch depends upon the mutual participation of all objectives, which can be mathematically given by the following equation:

$$F_{combined} = w_1 F_c + w_2 \sigma F_e, \quad w_2 = 1 - w_1 \tag{5}$$

w_1 represents the priority given to the cost function and w_2 represents the priority of the emission function. In this case, $w_1 = 1$ represents minimum fuel cost ($/hr) and $w_1 = 0$ represents the minimum emission (ton/hr), penalty factor σ($/ton) should be provided.

2.2 Non-dominated Approach

In multi-objective optimization, there is no unique solution but set of optimal solutions. Due to the conflicting nature of multiple objectives present, it is not possible for any single solution to satisfy the diverse requirements. Many researchers over the years have attempted to find various procedures for retaining solutions from non-dominated approach which are found while solving optimization problems. For example, a fast non-dominated sorting approach is represented in [12], which is capable of finding better spread of results, and more accurate solutions near the Pareto-optimal front. In [13], authors demonstrate a strength of Pareto approach for multi-objective optimization and use a mechanism of storing non-dominated solutions externally in a secondary continuously updated population. Literature [14] proposes a diversification generation method to randomly

generate trial solutions and reference set update method to renew the set of reference solutions with best ones. In our formulation, we have used an external array method to store the non-dominated solutions obtained in Pareto front.

2.3 Best Compromise Solution

While solving a multi-objective problem formulation, a set of optimal solutions are obtained. There should thus be a proper mechanism in place, to select a best compromise solution, which should reflect a trade-off between two objectives. So, after obtaining the Pareto front, decision-maker should choose one trade-off solution depending on the nature of specific problem formulation. A fuzzy membership function is developed given by the following equation:

$$\mu_{i,j} = \begin{cases} 1, f_j(X_i) \leq f_j^{\min} \\ \frac{f_j^{\max} - f_j(X_i)}{f_j^{\max} - f_j^{\min}}, f_j^{\min} \leq f_j(X_i) \leq f_j^{\max} \\ 0, f_j(X_i) \geq f_j^{\max} \end{cases} \tag{6}$$

The member function reflects the objective function's achievement while searching for optimal solution. In above equation, f_j^{min} gives value of objective function, highly satisfactory to decision-maker in respect to given problem ($\mu = 1$), whereas f_j^{max} gives value of objective function, highly unsatisfactory ($\mu = 0$). The normalized membership function is computed according to the following equation:

$$\mu_{i,j} = \frac{\sum_{j=1}^{k} \mu_{ij}}{\sum_{i=1}^{m} \sum_{j=1}^{k} \mu_{ij}}. \tag{7}$$

3 Particle Swarm Optimization (PSO)

PSO is a meta-heuristic evolutionary computation method, used to tackle complex optimization problems. It was first elaborated in Eberhart, Kennedy's paper in 1995. This algorithm has been developed from inspiration taken from bird and fish flocks in nature. In this work, time-varying acceleration coefficients (TVAC) are incorporated in particle swarm optimization [15] with inertia weight factor changing with iteration.

3.1 TVAC-PSO Components

Important components of TVAC-PSO method are illustrated below [16].

(1) *Particles in Optimization Algorithm*, X(t): Representation of obtained solutions, mathematically denoted by m vector, where m is total number of parameters to be optimized.
(2) *Swarm*: It is an apparently disorganized population of moving particles that tend to cluster together while each particle seems to be moving in a random direction.
(3) *Particle velocity*: The trajectory of each individual particle in *d*-dimensional search space is adjusted by adjusting the velocity of each particle. The position vector and velocity vector in the above-mentioned searching domain are $X_i = (x_{i1}, x_{i2}, x_{i3}, \ldots, x_{id})$ and $V_i = (v_{i1}, v_{i2}, v_{i3}, \ldots, v_{id})$, respectively. The best fitness value obtained by each particle at any given iteration can be illustrated as $P_i = (p_{i1}, p_{i2}, p_{i3}, \ldots.., p_{id})$, which is termed *pbest* and the fittest particle found so far is termed as *gbest* given by $P_g = (p_{g1}, p_{g2}, p_{g3}, \ldots\ldots, p_{gd})$. The new updated velocities and positions of the particles for next evaluation can be calculated using the following equations:

$$v_{id(iter)} = w_{(iter)} v_{id(iter-1)} + c_1 r_1 \left(p_{id(iter-1)} - x_{id(iter-1)}\right) + c_2 r_2 \left(p_{gd(iter-1)} - x_{id(iter-1)}\right) \tag{8}$$

$$x_{id(iter)} = x_{id(iter-1)} + v_{id(iter)}.$$

The coefficients of acceleration are c_1 and c_2 that are kept constant for classical PSO. For improved PSO considering TVAC, c_1 and c_2 are duly modified. The acceleration coefficient c_1 can be described as follows:

$$c_1 = c_{1f} - c_{1i} \times \frac{iter}{iter_{\max}} + c_{1i}. \tag{9}$$

Similarly, the acceleration coefficient c_2 can be described as follows:

$$c_2 = c_{2f} - c_{2i} \times \frac{iter}{iter_{\max}} + c_{2i}. \tag{10}$$

Here, c_{1i} and c_{2i} represent the initial values, whereas c_{1f} and c_{2f} represent the end numerical of self-searching and searching from other's experience. Generally, range in values of c_{1i} and c_{1f} can be chosen as 2.5 and 0.5, and the range of values of c_{2i} and c_{2f} can be chosen as 0.5 and 2.5, respectively [15].

(4) *Varying Inertia weight component*: This represents the parameter utilized for controlling influence of previous velocities on presently obtained velocities. So this reflects balance between global exploration and local exploration capabilities of solutions in search space. In starting process of searching, high inertia weight is used for increasing large range exploration among available solutions but then gradually lowered for enhancing local exploration. Generally inertia weight is regressive in nature from 0.9 to 0.4 and given as follows.

$$w = w_{\max} - \frac{w_{\max} - w_{\min}}{iter_{\max}} \times iter \tag{11}$$

$iter_{\max}$ represents total iterations and *iter* represents present iteration count.

(5) *Individual Best and Global Best*: As the particle navigates in searching domain, it measures its suitability at present position to foremost suitability it has attained at any time up to present time. The foremost position which is related to foremost fitness value tackled up till present is termed as individual best. The foremost position within total individual best positions achieved till present time is termed as global best [16].
(6) *Stopping criteria*: These are conditions for which searching will conclude. Here searching process concludes when iterations reach maximum limit.

4 Case Study and Results

The proposed technique is applied in IEEE 30 bus test case system [2]. The generator data for cost coefficients are shown in Table 1. The generation cost and emission levels using different weight factors are represented in Table 2 and best solution EED in IEEE 30 bus test system is represented in Table 3 (Figs. 1, 2 and 3).

Table 1 Cost coefficients of IEEE 30 bus test system

Bus no.	a_i	b_i	c_i	e_i	f_i
1	0.0	2.00	0.00375	50	0.063
2	0.0	1.75	0.01750	40	0.098
5	0.0	1.00	0.06250	0	0.000
8	0.0	3.25	0.00834	0	0.000
11	0.0	3.00	0.02500	0	0.000
13	0.0	3.00	0.02500	0	0.000

Table 2 Time-varying acceleration coefficients (TVAC-PSO) to find solution of economic emission dispatch problem for IEEE 30 bus test system

Lossy			Lossless		
W_1	Generation cost ($/hr)	Total emissions (tons/hr)	W_1	Generation cost ($/hr)	Total emissions (tons/hr)
0.0	654.0038	0.0785	0.0	646.1292	0.0737
0.1	651.0830	0.0799	0.1	636.2933	0.0752
0.2	647.2297	0.0881	0.2	630.6106	0.0816
0.3	645.5225	0.0892	0.3	626.4842	0.0817
0.4	636.2448	0.0925	0.4	622.5131	0.0995
0.5	627.9677	0.0982	0.5	616.8864	0.1393
0.6	619.4314	0.0998	0.6	603.4911	0.1447
0.7	613.3169	0.1672	0.7	602.1779	0.1536
0.8	602.7074	0.1680	0.8	600.8816	0.1586
0.9	594.8790	0.1702	0.9	592.6903	0.1613
1.0	592.3524	0.1798	1.0	589.9094	0.1694

Table 3 Best solution of economic emission dispatch problem for IEEE 30 bus test system

Objectives	Transmission network	Generation costs ($/h)	Emissions (tons/h)
Generation cost minimization	Lossless	589.9094	0.1694
	Lossy	592.3524	0.1798
Emission minimization	Lossless	646.1292	0.0737
	Lossy	654.0038	0.0785

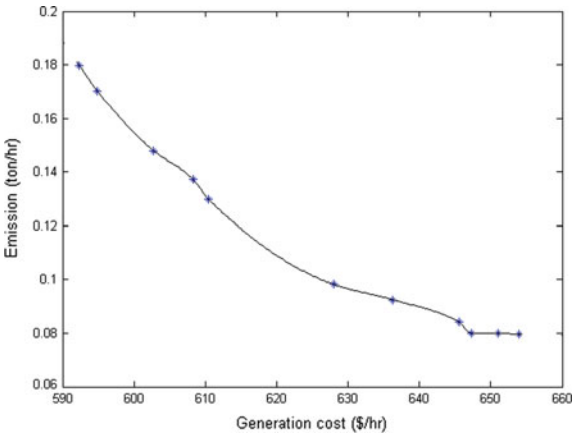


Fig. 1 Pareto front in IEEE 30 bus test case for lossy transmission system

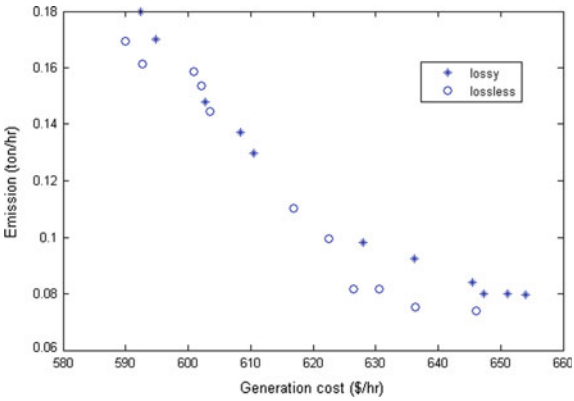


Fig. 2 Pareto front in IEEE 30 bus test case for lossy & lossless transmission system

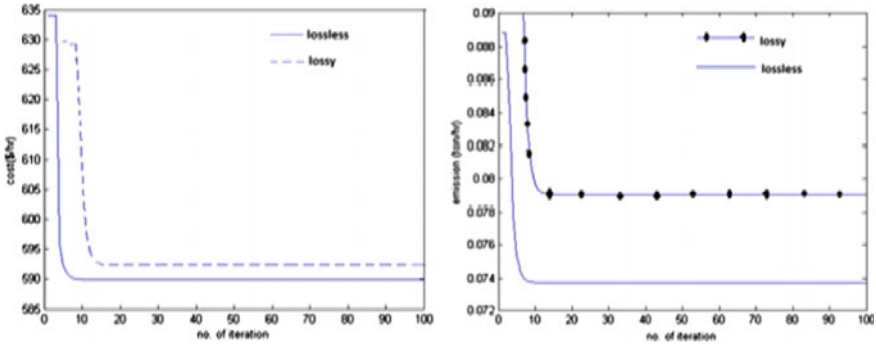


Fig. 3 Convergence characteristics for generation cost minimization and emission minimization for lossless & lossy transmission system

5 Conclusion

In this work, improved particle swarm optimization including time-varying acceleration coefficients is used to solve EED problem. Results obtained in economic emission dispatch problem by two methods discussed show that the latter help to obtain good solutions and superior Pareto front. The results represent high capability of TVAC-PSO technique for obtaining solution to EED problem.

Acknowledgements The authors express thanks to all members of Power Engineering department, Jadavpur University for providing support for research work. This research work is helped by project of (DRS and UPE II, UGC, DST-PURSE II, GOI) awarded to the Power Engineering department, Jadavpur University, Kolkata.

References

1. Mandal, K.K., Mandal, S., Bhattacharya, B., Chakraborty, N.: Non-convex emission constrained economic dispatch using a new self-adaptive particle swarm optimization technique. Appl. Soft Comput. **28**, 188–195 (2015)
2. Delshad, M.M., Rahim, N.A.: Multi-objective backtracking search algorithm for economic emission dispatch problem. Appl. Soft Comput. **40**, 479–494 (2016)
3. Madrigal, M., Quintana, V.H.: An analytical solution to the economic dispatch problem. IEEE Power Eng. Rev. **20**, 52–55 (2000)
4. Ji-Yuan, F., Lan, Z.: Real-time economic dispatch with line flow and emission constraints using quadratic programming. IEEE Trans. Power Syst. **13**, 320–325 (1998)
5. Farag, A., Al-Baiyat, S., Cheng, T.C.: Economic load dispatch multi-objective optimization procedures using linear programming techniques. IEEE Trans. Power Syst. 731–738 (1995)
6. Zwe-Lee, G.: Particle swarm optimization to solving the economic dispatch considering the generator constraints. IEEE Trans. Power Syst. **18**, 1187–1195 (2003)
7. Liang, Y.C., Cuevas Juarez, J.R.: A normalization method for solving the combined economic and emission dispatch problem with meta-heuristic algorithms. Int. J. Electr. Power Energy Syst. **54**, 163–186 (2014)
8. Wang, L., Singh, C.: Environmental/economic power dispatch using a fuzzified multi-objective particle swarm optimization algorithm. Electr. Power Syst. Res. 1654–1664 (2007)
9. Bhattacharya, A., Chattopadhyay, P.K.: Solving economic emission load dispatch problems using hybrid differential evolution. Appl. Soft Comput. **11**, 2526–2537 (2011)
10. Mousa, A.A.A.: Hybrid ant optimization system for multiobjective economic emission load dispatch problem under fuzziness. Swarm Evol. Comput. **18**, 11–21 (2014)
11. Jayakumar, D., Venkatesh, P.: Glowworm swarm optimization algorithm for solving multiple objective environmental economic dispatch problem. Appl. Soft Comput. 375–386 (2014)
12. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic Algorithm:NSGA-II. IEEE Trans. Evol. Comput. **6**, 182–197 (2002)
13. Zitzler, E., Thiele, L.: Multi objective evolutionary algorithms: a comparative case study and the strength pareto approach. IEEE Trans. Evol. Comput. **3**(4), 257–271 (1999)
14. Silva, C.E., de Athayde, M., Klein, C.E., Mariani, V.C., dos Santos Coelho, L.: Multiobjective scatter search approach with new combination scheme applied to solve environmental/economic dispatch problem. Energy, vol. 53, pp. 14–21 (2013)
15. Ratnaweera, A., Halgamuge, S.K., Watson, H.C.: Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients. IEEE Trans. Evol. Comput. **8**(3), 240–255 (2004)
16. Abido, M.A.: Optimal power flow using particle swarm optimization. Electr. Power Energy Syst. **24**, 563–571 (2002)

Design of Rule-Based Load Frequency Controller for Multi-machine System

Jyotirmoy Biswas and Parthasarathi Bera

Abstract This paper presents an application of rule-based proportional integral (PI) controller for the load frequency control (LFC) taking into account the effect of power transmission network. The analysis is carried out considering a six-bus system and genetic algorithm (GA) is used to optimize the gains of conventional as well as rule-based PI controller. The results show that dynamic response of frequency deviation improved significantly using rule-based PI controller compared to conventional PI controller.

Keywords Load frequency control (LFC) · Genetic algorithm (GA) · Rule-based PI controller

1 Introduction

In electrical power systems both electrical power generations and loads vary continuously. There will be change in frequency if load changes immediately and as a result there will be change in net power interchanges with the connected control area [1, 2].

Load frequency control is very important for designing electrical power system and by the time various control strategies have been proposed [3–7]. Tripathy et al. [4] have designed load frequency controller for power system with reheat steam turbines and governor deadband nonlinearity. Hiyama [5] has designed load frequency controller for decentralized interconnected power system. Lim et al. [6] have designed decentralized load frequency controller for multi-area power system.

J. Biswas (✉) · P. Bera
Department of Electrical Engineering, Kalyani Government Engineering College, Kalyani 741235, India
e-mail: jyotirmoy@gmail.com

P. Bera
e-mail: parthabera1977@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_6

Yang et al. [7] have designed decentralized load frequency controller based on structured singular values.

In the present model of control area, where a number of generating units are represented by equivalent one turbine generator system is used to design the controller and theb effect of transmission of power through transmission line is ignored and the loads on distributed load buses are summed up and are considered to be applied on a particular generator. The weak tie lines between different control areas have been separately incorporated for design and analysis of load frequency control [1, 4–7].

When the control of the traditional vertically integrated utilities has been considered by this reduced order model, it works efficiently. But due to recent development of the spot market of electricity, utilities have no direct control in their control area. This new development, which is known as restructured power system, requires new control schemes and strategies [8–10]. Iracleous et al. [11] have designed the controller for multi-task automatic generation control for power regulation.

In the present condition, the existing LFC model is not technically suitable and for designing the controller and each generating station should be considered separately. The interconnection between generators and load buses cannot be considered as a single electrical net, so the concept of considering individual power station as a control area is not a correct approach. As a result proposition of a new model for control area is a very important to design load frequency controller. Biswas et al. [12] have considered a detailed model of LFC for practical electrical network considering the effect of power transmission network and used GA for optimizing the gains of PI controller for generators of a six-bus power system for improving the dynamic response of the system. In the present work, rule-based PI controller has been designed to improve the dynamic response of the system and the rules have been set based on frequency deviation and rate of change of frequency deviation for each individual generator.

2 Development of System Model

Figure 1 shows the block diagram of LFC with rule-based PI controller of a single generator and the state space model for each generator is given as [13].

$$\dot{x}_i = \begin{bmatrix} -\frac{1}{T_{p_i}} & \frac{K_{p_i}}{T_{p_i}} & 0 & 0 \\ 0 & -\frac{1}{T_{r_i}} & \left(\frac{1}{T_{r_i}} - \frac{K_{r_i}}{T_{t_i}}\right) & \frac{K_{r_i}}{T_{t_i}} \\ 0 & 0 & -\frac{1}{T_{t_i}} & \frac{1}{T_{t_i}} \\ -\frac{1}{R_i T_{g_i}} & 0 & 0 & -\frac{1}{T_{g_i}} \end{bmatrix} x_i + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{T_{g_i}} \end{bmatrix} U_i - \begin{bmatrix} \frac{K_{p_i}}{T_{p_i}} \\ 0 \\ 0 \\ 0 \end{bmatrix} \Delta P_{g_i}, \qquad (1)$$

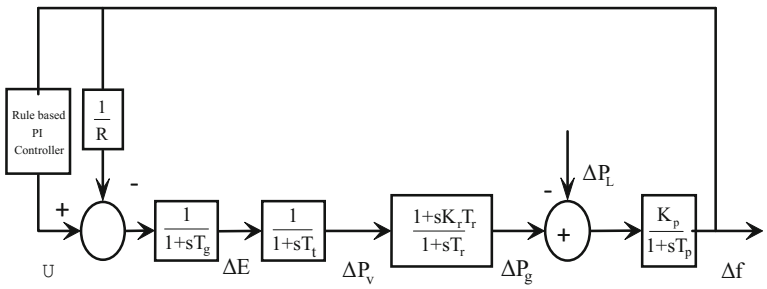


Fig. 1 The block diagram of single area LFC [13]

where $x_i = \left[\Delta f_{g_i}\ \Delta P_{g_i}\ \Delta P_{v_i}\ \Delta E_i\right]^T$ of ith generator; and T_p, T_t, T_g, R, K_p, K_r are the physical constants of the system. Here inputs to the PI controllers are frequency deviations of the alternators.

The above Eq. (1) can be written as

$$\dot{x}_i = A_i x_i + B_i U_i - G_i \Delta P_{g_i}. \tag{2}$$

Considering a system having 'm' number of generators, the equation for a particular generator can be written as follows [11]:

$$\dot{x} = Ax + BU - G\Delta P_g. \tag{3}$$

Here,

$x = [x_1\ x_2, \ldots . x_m]^T$, $U = [U_1\ U_2, \ldots, U_m]^T$ and $\Delta P_g = [\Delta P_{g1}\ \Delta P_{g2}, \ldots, \Delta P_{gm}]$.

The equation of real and reactive power injected at each bus of any N bus power system can be written as follows [11]:

$$P_k = |V_k| \sum_{j=1}^{N} ((G_{kj}\cos\delta_{kj} + B_{kj}\sin\delta_{kj})|V_k|) \tag{4}$$

$$Q_k = |V_k| \sum_{j=1}^{N} ((G_{kj}\sin\delta_{kj} - B_{kj}\cos\delta_{kj})|V_k|). \tag{5}$$

After linearization, these equations can be written in the following matrix form:

$$\begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} = \begin{bmatrix} H & N \\ M & L \end{bmatrix} \begin{bmatrix} \Delta\delta \\ \frac{\Delta|V|}{|V|} \end{bmatrix}, \tag{6}$$

where

$$\begin{aligned} H_{kj} = L_{kj} &= |V_k||V_j|(G_{kj}\sin\delta_{kj} - B_{kj}\cos\delta_{kj}) \\ &= |V_k||V_j|(G_{kj}\delta_{kj} - B_{kj}) \end{aligned} \tag{7}$$

and

$$H_{kk} = -Q_k - B_{kk}V_k^2 \quad \text{and} \quad L_{kk} = Q_k - B_{kk}V_k^2. \tag{8}$$

As real power changes are less sensitive to voltage magnitude and reactive power changes are less sensitive to angle, we can take

$N = 0$ and $M = 0$,

Therefore, from Eq. (6) we get

$$\Delta P = \overline{H}\Delta\overline{\delta} \tag{9}$$

ΔP can be expressed as follows:

$$\Delta P = \begin{bmatrix} \Delta P_g \\ \Delta P_l \end{bmatrix} = \begin{bmatrix} \overline{H}_{gg} & \overline{H}_{gl} \\ \overline{H}_{lg} & \overline{H}_{ll} \end{bmatrix} \begin{bmatrix} \Delta\overline{\delta}_g \\ \Delta\overline{\delta}_l \end{bmatrix}. \tag{10}$$

Here, $\Delta\overline{\delta}_g$ and $\Delta\overline{\delta}_l$ are the relative angle differences at the generator and load buses, respectively.

So from Eq. (10) we have

$$\Delta\overline{\delta}_l = -\overline{H}_{ll}^{-1}\overline{H}\Delta\overline{\delta}_g + \overline{H}_{ll}^{-1}\Delta P_l. \tag{11}$$

Putting the above expression of $\Delta\overline{\delta}_l$ in the first row of Eq. (10), we obtain

$$\Delta P_g = M_1\Delta\overline{\delta}_g + M_2\Delta P_l, \tag{12}$$

where

$$M_1 = \overline{H}_{gg} - \overline{H}_{gl}\overline{H}_{ll}^{-1}\overline{H}_{lg} \quad \text{and} \quad M_2 = \overline{H}_{gl}\overline{H}_{ll}^{-1}. \tag{13}$$

Putting the expression of ΔP_g from Eq. (12) in Eq. (3) we get the system as [11]

$$\begin{bmatrix} \dot{x} \\ \Delta\delta_g \end{bmatrix} = \begin{bmatrix} A & -GM_1 \\ \alpha & 0 \end{bmatrix} \begin{bmatrix} x \\ \Delta\delta_g \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u - \begin{bmatrix} GM_2 \\ 0 \end{bmatrix} \Delta P_l, \tag{14}$$

where $\alpha = [E, \text{diag}(\alpha_{i+1})]$, $\alpha_i = [0\,0\,2\pi]$, E is the $(m-1)\times 3$ matrix: $E = [-\alpha_1^T, \ldots, -\alpha_1^T]^T$. The state space model of LFC for each generating station is represented by Eq. (14).

3 Objective Function

For optimizing the gains of conventional and rule-based PI controllers using genetic algorithm (GA), the following objective function based on the integral of time-multiplied square error (ITSE) criterion has been used [14, 15]:

$$J = \int_0^\infty t(\Delta f)^2 dt. \tag{15}$$

In the present problem, a step disturbance in load change in the load bus ($\Delta P_l = 0.2$ pu) has been used to compute the value of J for optimizing the gains of both rule-based and PI controllers.

4 Genetic Algorithm

John Holland in his book in Natural and Artificial systems [16] first suggested genetic algorithms (GAs) to search randomly the best answer for optimization problems. In the present problem, nonlinear multivariable functions are to be optimized to tune the gains of PI controllers and the values of crossover probability $P_c = 1.0$, mutation probability $P_m = 0.006$ have been considered for tuning the gains of rule-based and PI controllers.

5 Proposed Rule-Based Gain Scheduling Controller

In this present work, change of frequency deviation (Δf) and rate of change of frequency deviation $(\Delta \dot{f})$ have been considered to schedule the gains of PI controller and while optimizing the gains using GA, the values of K_P and K_I are assumed in the range $[K_{P\ min}\ K_{P\ max}]$ and $[K_{I\ min}\ K_{I\ max}]$, respectively. The parameters of K_P and K_I are determined by the following set of rules:

If $\Delta f(k)$ is A_i and $\Delta \dot{f}(k)$ is B_i then K_P is C_i and K_I is D_i
for i = 1, 2, …..n.

In this case, four rules have been considered for setting the gains of PI controller for each generator using the positive and negative values of change of frequency deviation (Δf) and rate of change of frequency deviation $(\Delta \dot{f})$, respectively.

6 Result

In the present work, a six-bus system is considered as shown in Fig. 2 and the bus and line data for that system are given in **Appendix**. The proportional integral controller is designed for the generators by minimizing the objection function as given by Eq. (18) using genetic algorithm and the range for optimizing the gains of proportion and integral controller has been considered $K_{Pmax} = 30$, $K_{Pmin} = -30$ and $K_{Imax} = 25$, $K_{Imin} = -25$, respectively. The optimized values of gains of PI controllers are given as $K_{P1} = 14.9089$, $K_{P2} = 17.9738$ and $K_{I1} = -5.1461$, $K_{I2} = 9.9224$. Here as the optimization problem is nonlinear in nature, the optimum value of K_{I1} has been found negative. For rule-based controller, again the values of K_P and K_I have been optimized under the condition that each rule has been satisfied while optimizing the objective function as given by Eq. (18) and the optimized values of K_P and K_I are given in Tables 1, 2, and 3 for generator 1, 2, and 3, respectively.

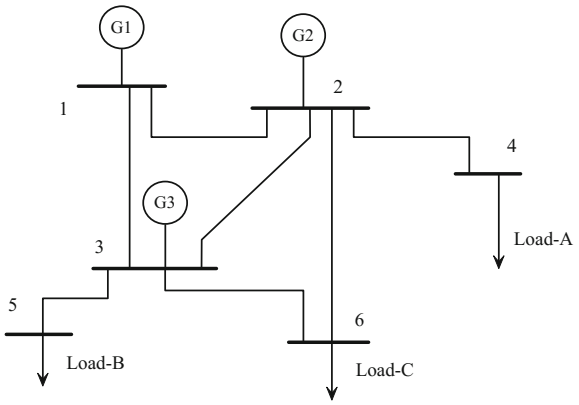


Fig. 2 The six-bus power system [12]

Table 1 Value of gains of rule-based PI controller for Generator-1

$\dot{\Delta f}$ / Δf	+ Ve	−Ve
+Ve	$K_I = -7.5$ $K_P = 17.4$	$K_I = -8$ $K_P = 15.9$
−Ve	$K_I = 8$ $K_P = 16.8$	$K_I = 5$ $K_P = 14.9$

Table 2 Value of gains of rule-based PI controller for Generator-2

$\dot{\Delta f}$ / Δf	+ Ve	−Ve
+Ve	$K_I = -6.5$ $K_P = 10.9$	$K_I = -5$ $K_P = 12.9$
−Ve	$K_I = 5$ $K_P = 16.9$	$K_I = 2$ $K_P = 14.9$

Table 3 Value of gains of rule-based PI controller for Generator-3

$\Delta \dot{f}$ / Δf	+ Ve	−Ve
+Ve	$K_I = 7.9$ $K_P = 19.5$	$K_I = 8.9$ $K_P = 20.9$
−Ve	$K_I = 10.9$ $K_P = 18.9$	$K_I = 9.9$ $K_P = 17.9$

Figures 3 and 4 show the dynamic response of frequency of Generator-1 for 20% load change in load bus 5 and 6, respectively. Figures 5 and 6 show the dynamic responses of frequency deviation of Generator-2 with conventional PI controller and with rule-based PI controller for 20% change of load of the load buses 4 and 5, respectively. Figures 7 and 8 show the dynamic responses of frequency deviation of Generator-3 with conventional PI controller and with rule-based PI controller for 20% change of load of the load buses 4 and 5, respectively. From these figures, it is seen that the peak overshoot and settling time of dynamic responses of frequency

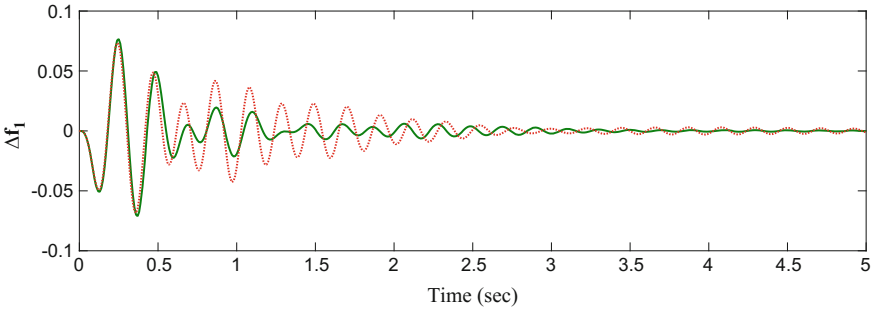


Fig. 3 Dynamic response of Generator-1 considering conventional PI control and rule-based control for20% load change in bus 5 (————with rule-based PI control, - - - - - - -with conventional PI control)

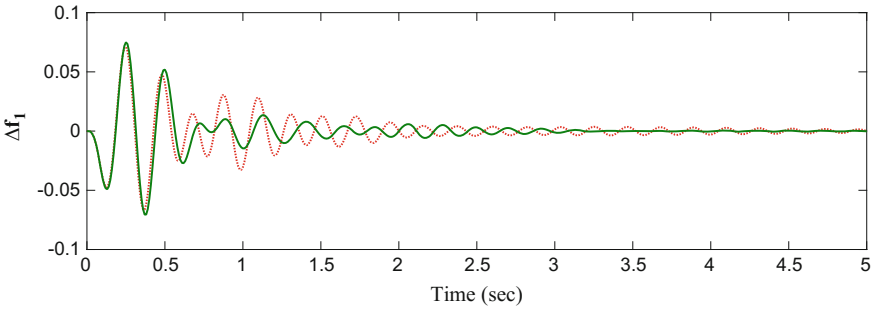


Fig. 4 Dynamic response of Generator-1 considering conventional PI control and rule-based control for 20% load change in bus 6 (————with rule-based PI control, - - - - - - -with conventional PI control)

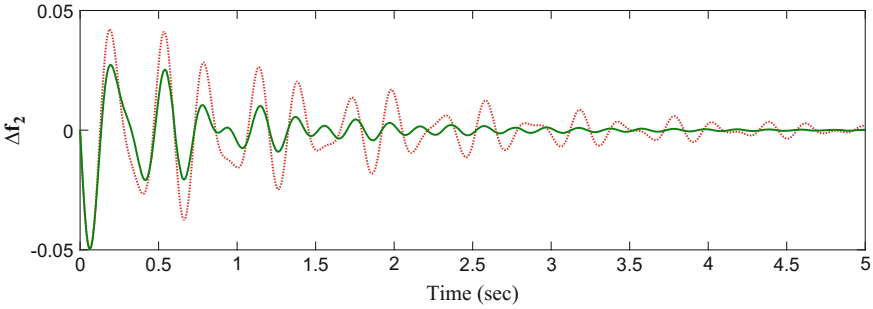


Fig. 5 Dynamic response of Generator-2 considering conventional PI control and rule-based control for 20% load change in bus 4 (————with rule-based PI control, - - - - - - - with conventional PI control)

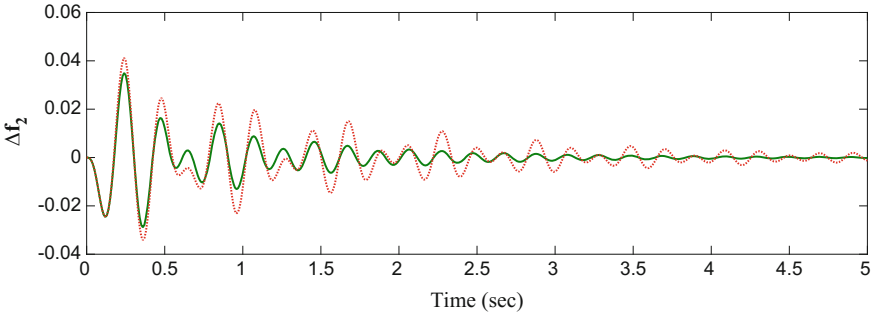


Fig. 6 Dynamic response of Generator-2 considering conventional PI control and rule-based control for 20% load change in bus 5 (————with rule-based PI control, - - - - - - - with conventional PI control)

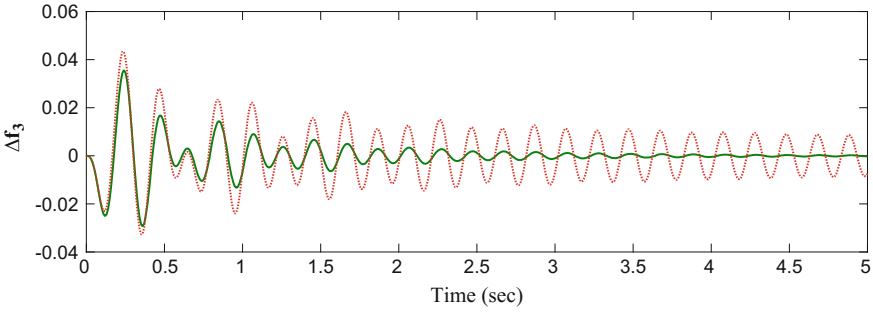


Fig. 7 Dynamic response of Generator-3 considering conventional PI control and rule-based control for 20% load change in bus 4 (————with rule-based PI control, - - - - - - - with conventional PI control)

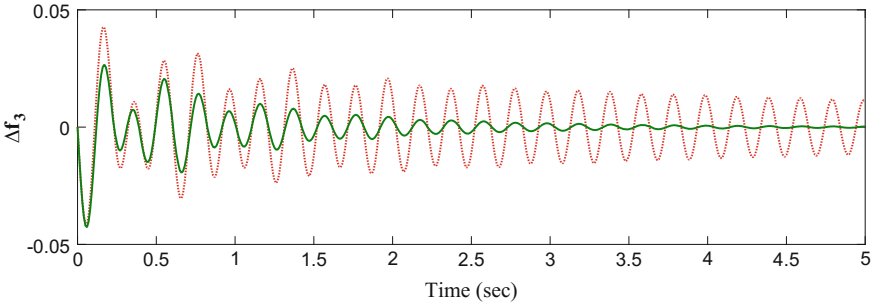


Fig. 8 Dynamic response of Generator-3 considering conventional PI control and rule-based control for 20% load change in bus 5 (————with rule-based PI control, - - - - - - -with conventional PI control)

deviation improve with rule-based PI controller compared to conventional PI controller.

7 Conclusion

In this paper, a detailed model of control area for load frequency control (LFC) has been considered to take into account the effect of power transmission network. Gains of conventional proportional integral (PI) controller and rule-based PI controller have been optimized using genetic algorithm (GA) for designing the load frequency controller of each generator of a particular six-bus system. The dynamic responses of frequency deviation of each generator have been observed with conventional PI controller and with rule-based PI controller for 20% change of load of different load buses. Simulation results show that the peak overshoot and settling time of dynamic responses for frequency deviation improve with rule- based PI controller compared to conventional PI controller.

Appendix

See Tables 4, 5, and 6.

Table 4 Machine data

	T_p	T_t	T_g	R	K_p	K_r
Gen-1	20.67	0.3	0.08	2.4	130	0.35
Gen-2	20	0.1	0.08	2.5	120	0.33
Gen-3	22	0.3	0.1	2.4	123	0.33

Table 5 Bus data

Bus code	Voltage (in p.u.)		Load (in p.u.)		Generation (in p.u.)			
	Magnitude	Phase angle	MW	MVAr	MW	MVAr	Q_{min}	Q_{max}
1	1.06	0.0	0	0	0	0	0	0
2	1.00	0.0	0	0	20	0	–30	30
3	1.00	0.0	0	0	50	0	–40	40
4	1.00	0.0	50	35	0	0	0	0
5	1.00	0.0	50	25	0	0	0	0
6	1.00	0.0	60	25	0	0	0	0

Table 6 Line data

Bus code	R (in p.u.)	X (in p.u.)	B/2 (in p.u.)
1–2	0.02	0.08	0.00
1–3	0.02	0.06	0.00
2–3	0.06	0.25	0.00
2–4	0.06	0.25	0.00
2–6	0.06	0.25	0.00
3–5	0.06	0.25	0.00
3–6	0.06	0.08	0.00

References

1. Elgerd, O.: Electric Energy Systems Theory: An Introduction. McGraw-Hill, New York (1971)
2. Jaleeli, N., VanSlyck, L., Ewart, D., Fink, L., Hoffman, A.: Understanding automatic generation control. IEEE Trans. Power Syst. **7**(3), 1106–1112 (1992)
3. Fosha, C., Elgerd, O.: The megawatt-frequency control problem: A new approach via optimal control theory. IEEE Trans. Power Appar. Syst. **89**, 563–577 (1970)
4. Tripathy, S.C., Hope, G.S., Malik, O.P.: Optimisation of load–frequency control parameters for power systems with reheat steam turbines and governor deadband nonlinearity. IEE Proc. **129**, pt. C, 10–16 (1982)
5. Hiyama, T.: Design of decentralised load–frequency regulators for interconnected power systems. IEE Proc. **129**, 17–23 (1982)
6. Lim, K.Y., Wang, Y., Zhou, R.: Robust decentralised load–frequency control of multi-area power systems. IEE Proc Gener. Transm. Distrib. **143**, Pt. C, 377–386 (1996)
7. Yang, T.C., Cimen, H., Zhu, Q.M.: Decentralised load–frequency controller design based on structured singular values. IEE Proc. Gener. Transm. Distrib. **145**, 7–14 (1998)
8. Bialek, J.: Tracing the flow of electricity. IEE Proc. Gener. Transm. Distrib. **143**, 313–320 (1996)
9. Christie, R.D., Bose, A.: Load frequency control issues in power system operation after deregulation. IEEE Trans. Power Syst. **11**(3), 1191–1196 (1996)
10. Zobian, A., Ilic, M.D.: Unbundling of transmission and ancillary services-technical issues. IEEE Trans. Power Syst. **12**, 539–548 (1997)
11. Iracleous, D.P., Alexandridis, A.P.: A multi-task automatic generation control for power regulation. Electr. Power Syst. Res. **73**, 275–285 (2005)

12. Biswas, J., Bera, P.: PI-Based load frequency controller design for multi machine system using genetic algorithm. Reason.-A Tech. J. **XIV**(7), 65–76 (2015)
13. Das, D.: Electrical Power Systems, A Book. New Age International (P) limited (2006)
14. Schultz, W.C., Rideout, V.C.: Control system performance measures: past, present and future. IRE Trans. Autom. Control, AC-6 **22**, 22–35 (1961)
15. Ogata, K.: Modern Control Engineering, A Book, pp. 293–313. Printice-Hall, Englewood cliffs, NJ (1970)
16. Holland, J.H.: Adaptation in Nature and Artificial Systems. University of Michigan Pres, Am Arrbor (1975)
17. Biswas, S., Bera, P.: GA application to optimization of AGC in two–area power system using battery energy storage. In: International Conference on Communications, Devices and Intelligent Systems (CODIS), Jadavpur University (2012)
18. Debbarma, S., Saikia. L.C., Sinha, N.: AGC of a multi-area thermal system under deregulated environment using non-integer controller. Electr. Power Syst. Res. **95**, 175–183 (2013)
19. Sahu, R.K., Panda, S., Rout, U.K., Sahoo, D.K.: Teaching learning based optimization algorithm for automatic generation control of power system using 2-DOF PID controller. Int. J. Electr. Power Energy Syst. **77**, 287–301, May 2016
20. Madichetty, S., Dasgupta, A., Kumar, L.V.S.: Application of modular multi level converter for AGC in interconnected power system. Int. J. Electr. Power Energy Syst. **74**, 293–300 (2016)

Soft Computing Approach to Electrical Transmission Network Congestion Management

Debapriya Sur Mukhopadhyay, Reshmi Chanda, Debjani Chakraborti and Papun Biswas

Abstract In this paper an efficient technique is described for managing the congestion in electric transmission network based on rescheduling the nearby generators and/or shedding some of the loads. To incorporate the uncertainty in the system objectives and parameters, fuzzy environment is considered for the formulation of the problem. In the solution process bio-inspired computational technique, genetic algorithm (GA) is used. The approach is illustrated by standard IEEE 30-bus 6-generator test system.

Keywords Congestion management · Fuzzy set · Fuzzy programming · Fuzzy goal programming · Genetic algorithm · Load shedding

1 Introduction

In a power supply system, when a transmission line hits its transmission limits, the state is referred to congestion. Actually, in a interconnected transmission network, the reliably of smooth transfer of electric power depends on various characteristics of power generation and dispatch to demand centers, where voltage drop and line loading are the key factors for occurrence of any hazard in power supply system.

D.S. Mukhopadhyay (✉)
Chandernagore Municipal Corporation, Chandannagar, WB, India
e-mail: mukhopadhyay.debapriya@gmail.com

R. Chanda
Abacus Institute of Engineering and Management, Mogra, WB, India
e-mail: c.reshmi84@gmail.com

D. Chakraborti
Narula Institute of Technology, Agarpara, WB, India
e-mail: debjani.chakraborti@gmail.com

P. Biswas (✉)
JIS College of Engineering, Kalyani, WB, India
e-mail: papunbiswas@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_7

Congestion [1] in transmission lines generally occurs due to generator shutdown, increase in load, or failure of some setup components of the system.

The optimization method for voltage drop estimation and the line loading and thereby smoothing power flow was first studied by Abiad and Stagg (1963) in [2]. Thereafter, different classical load flow-based optimization models [3, 4] for CM have been developed.

The decomposition of spot prices technique to estimate congestion cost component was presented by Finney et al. (1997) in [5]. The aspect of congestion management (CM) in a real-time operational environment was discussed in [6, 7]. The study on congestion minimization problem based on AC-OPF method was proposed by Rau (2000) [8]. Further, utilization of thyristor-controlled series compensation (TCSC) in [9] for reduction of congestion cost was proposed by Lee (2002) in [10]. Again, in the context of solving congestion problem, Rodrigues and DaSilva (2003) in [11] proposed load curtailment criterion to minimize congestion hazard. The use of location marginal price method to OPF-based transmission congestion was studied by Milano et al. (2003) in [12]. An effective congestion management technique with rescheduling of real and reactive powers concerning optimal allocation of reactive power resources was studied by Kumar et al. (2004) [13]. An economic method for management of congestion was proposed by Talukdar et al. (2005) in [14].

Again, different heuristic methods [15] have been addressed in [16–18] to reach optimal decisions of CM problems in uncertain environment. The fuzzy-based soft computing approach for collapse sequences arise with CM has also been studied by Hazra and Sinha (2009) in [19]. Fuzzy adaptive bacterial foraging algorithm in [20] and fuzzy adaptive gravitational search algorithm for CM have been proposed by Venkaiah and Vinod Kumar (2011) in [21] and Vijaya Kumar et al. (2013) in [22], respectively.

In this article, a fuzzy goal programming (FGP) technique based on the importance of the objectives is proposed for optimal CM decision for simultaneous optimization of operational cost, overload alleviation, and power loss in transmission lines under a set of system constraints in decision environment. In the proposed approach, a GA technique is used in model formulation and then performing model simulations to arrive at the optimal CM decision. The technique is tested on IEEE 30-bus test system. The solution is compared with the approach studied in [16] previously to illustrate the effectiveness of the approach.

Now, formulation of a CM problem is described in Sect. 2.

2 Problem Statement

The three major objective functions that are in general involved with a CM problem are discussed in the following section.

2.1 Problem Objectives

- **Alleviation of overload**

Transmission line overload alleviation is an essential power system practice to ensure secured and stable system operation and prevent the occurrence of system outage. Transmission line overload can be alleviated by lines switching, generation rescheduling, and load shedding.

The overload alleviation function is appeared as follows:

$$F_1 = \sum_{i=1}^{nl} (S_i - S_i^{\max})^2, \tag{1}$$

where F_1 is the cumulative overload, nl is the number of overloaded lines, and S_i *and* $S_i^{\max}$ are the mega volt ampere (MVA) flow and MVA capacity of line i, respectively.

- **Operational cost**

The total operational cost involved with CM can be expressed as the sum of fuel cost and the cost of load shedding.

The total operation cost function can be expressed as

$$F_2 = \sum_{i=1}^{ng} (a_i + b_i P_{gi} + c_i P_{gi}^2) + \sum_{k=1}^{pl} (a_k' + b_k' L_{shd,k} + c_k' L_{shd,k}^2), \tag{2}$$

where F_2 is the total operating cost, ng is the number of participating generators, pl represents the number of participating loads, P_{gi} is the power generation from ith generator, $L_{shd,k}$ is the amount of load shedding at bus k, and a_i, b_i, c_i are the cost coefficients associated with generation of power from generator g_i, and a_k', b_k', c_k' are the cost coefficients of load shedding at bus k.

- **Power-loss function**

The function associated with power transmission lines that directly affect the ability to transfer power effectively is the power-loss function. The real power loss (F_3) in megawatt (MW) can be defined as

$$F_3 = \sum_{l=1}^{L} g_l [V_i^2 + V_j^2 - 2V_i V_j \cos(\delta_i - \delta_j)], \tag{3}$$

where 'L' is the number of transmission lines, g_l is the conductance of the lth line, V_i and V_j are the voltage magnitude, δ_i and δ_j are the voltage phase angle at the end buses i and j of the lth line, respectively, and '*cos*' stands for *cosine* function.

2.2 Objective Constraints

The system constraints that are inherently associated with power generation system are defined as follows.

- **Power balance constraint**

The power balance constraint can be expressed as

$$\begin{aligned} P_{gi} - P_{di} - V_i \sum_{j=1}^{H} V_j [G_{ij} \cos\ (\delta_i - \delta_j) + B_{ij} \sin\ (\delta_i - \delta_j)] = 0 \\ Q_{gi} - Q_{di} - V_i \sum_{j=1}^{H} V_j [G_{ij} \sin\ (\delta_i - \delta_j) + B_{ij} \cos\ (\delta_i - \delta_j)] = 0 \end{aligned} \tag{4}$$

where $j = 1,2, \ldots, H$; and H is the number of total buses; real and reactive powers of the generator connected to the ith bus are represented by P_{gi} and Q_{gi}, respectively; the real and reactive powers of the load connected to the ith bus are represented by P_{di} and Q_{di}, respectively; the transfer conductance and susceptance between bus i and bus j are represented by G_{ij} and B_{ij}, respectively; and δ_i and δ_j are bus voltage angle of bus i and bus j, respectively.

- **Generator capacity constraint**

Following the conventional power generation and dispatch system, the constraints on the outputs of the generators can be considered as

$$\begin{aligned} P_{g_i}^{\min} \le P_{g_i} \le P_{g_i}^{\max}, \\ Q_{g_i}^{\min} \le Q_{g_i} \le Q_{g_i}^{\max}, \quad i = 1, 2, \ldots, N, \end{aligned} \tag{5}$$

where P_{gi} and Q_{gi} are real and reactive power outputs of the generators and where min and max stand for minimum and maximum values, respectively.

- **Voltage constraint**

In a power generation and dispatch system, the constraints on the generators' voltages can be considered as

$$V_i^{\min} \le V_i \le V_i^{\max};\ i = 1, 2, \ldots, N, \tag{6}$$

where *min* and *max* stand for minimum and maximum values, respectively.

3 Proposed Approach

3.1 Overview

The objectives of a DM are described fuzzily, whereas system constraints are taken crisp, which is common in contrast to general structure of power management problems. The fuzzy goals are characterized by their membership function. In problem formulation priority-based FGP approach is considered to achieve an optimal solution from the conflicting objectives.

In the process of designing the model of CM, it is to be observed that the objectives and most of the system constraints of the problem are nonlinear in nature. To avoid complexity in [23] in computational process with nonlinearity in goals/constraints and to avoid the load with the use of a traditional linear approximation approach in [24], a GA approach as a goal satisfier in [25] rather than objective optimizer for multiobjective decision analysis is introduced in the solution search process.

3.2 Compromise Solution

Due to the fuzzy nature of the objective functions, ith solution is represented by the membership function as in [26]:

$$\mu_j = \begin{cases} 1 & \text{if } F_j \le F_j^{\min}, \\ \dfrac{F_j^{\max} - F_j}{F_j^{\max} - F_j^{\min}} & \text{if } F_j^{\min} < F_k \le F_j^{\max}, \\ 0 & \text{if } F_j > F_j^{\max}, \end{cases} \tag{7}$$

Using the membership function, fuzzy goal is formulated as

$$\frac{F_j^{\max} - F_j}{F_j^{\max} - F_j^{\min}} + d_j^- - d_j^+ = 1, \tag{8}$$

where $d_j^-, d_j^+ \ (\ge 0)$ with $d_j^- . d_j^+ = 0$, represent the under- and over-deviational variables associated with the jth membership goal.

Now, the main objective is to reduce the under-deviation from all the fuzzy goals simultaneously and find out an optimal compromise solution. In this respect, the priority-based FGP model is formulated where the under-deviation is minimized as per the importance of the objective. Once the objective with first priority is considered, the solutions are locked and move to the next objective with second priority and so on.

Now, following the proposed procedure, construction of membership goals of the defined fuzzy goals of all the objectives of the problem and then formulation the priority-based executable model for power generation decision are implemented via an application presented in Sect. 4.

4 Application of the Proposed Approach

The IEEE 30-bus 6-generator test system in [14] is considered to demonstrate the potential use of the approach. The test system has 6 generators, 41 lines, and 30 buses. The total system has 21 load buses with 283.40 MW total demands.

The data associated with different types of cost coefficients and load shedding cost coefficients power generation, different transmission lines and loads at different buses are given in standard IEEE 30-bus 6-generator test system in [14].

The simulation runs conducted for execution of the problem are presented in Table 1.

The MATLAB-based optimization toolbox is used for the solution of the problem in Intel (R) Pentium CPU with 2.00 GHz clock pulse and 2 GB RAM. In the solution process following GA parameters are considered: Initial population = 100, Roulette wheel selection method, Single-point crossover with 0.8 crossover probability, mutation probability = 0.07, and maximum generation = 200.

Then, following the procedure and using the data presented, the membership goals of the defined fuzzy goals can be obtained by addressing the second goal expression in (3).

Then, the executable FGP models designed under a given priority structure for individual three simulation runs are presented as follows.

- **The model under Run-1**

The model for simulation of overload with reduction of capacity of line 1 from 130 to 50 MW is designed as follows.

Table 1 Simulation runs experiments

Sl. no.	Simulated cases
1	Reducing capacity of line 1–2 by 61.54% and thereby creating the overload condition
2	Reducing capacity of two transmission lines: line 1–3 by 61.54% and 2–4 by 76.92% and thereby creating a critical overload condition
3	Shutdown of unit 3 at bus 5 and reducing the capacity of line 2–5 by 61.54% and thereby creating another critical overload condition

Find $(S_1, P_{g_i}, i = 1, 2, 5, 8, 11, 13)$ so as to:

$$Minimize\, Z = \left[P_1 \left(\frac{1}{3.0 - 0.5} \right) d_1^-,\ P_2 \left(\frac{1}{557942 - 530000} \right) d_2^-,\ P_3 \left(\frac{1}{5.50 - 3.50} \right) d_3^- \right]$$

and satisfy

$$\mu_{F_1}: \frac{3.0 - (S_{1-2} - 50)^2}{3.0 - 0.5} + d_1^- - d_1^+ = 1,$$

$$\mu_{F_2}: \frac{\begin{array}{l} 547942 - (1650P_1 + 1.25P_1^2 + 1875P_2 + 1.425P_2^2 + 1875P_5 \\ + 1.425P_5^2 + 2025P_8 + 1.5P_8^2 + 2025P_{11} + 1.5P_{11}^2 + 2025P_{13} + 1.5P_{13}^2 \end{array}}{537942 - 530000} + d_2^- - d_2^+ = 1,$$

$$\mu_{F_3}: \frac{5.50 - P_L}{2.50} + d_3^- - d_3^+ = 1,$$

subject to

$$P_{G1} - V_1[V_2\{5.2246\cos(\delta_1 - \delta_2) - 15.6467\sin(\delta_1 - \delta_2)\} + V_3\{1.2437\cos(\delta_1 - \delta_3) - 5.0960\sin(\delta_1 - \delta_3)\}] = 0 \tag{9}$$

$$\begin{array}{l} P_{G2} - 0.217 - V_2[V_1\{5.2246\cos(\delta_2 - \delta_1) - 15.6467\sin(\delta_2 - \delta_1)\} \\ + V_4\{1.7055\cos(\delta_2 - \delta_4) - 5.1974\sin(\delta_2 - \delta_4) + V_5\{1.1360\cos(\delta_2 - \delta_5) \\ - 4.7725\sin(\delta_2 - \delta_5)\} + V_6\{1.6861\cos(\delta_2 - \delta_6) - 5.1165\sin(\delta_2 - \delta_6)\}] = 0 \end{array} \tag{10}$$

$$\begin{array}{l} P_{G5} - 0.942 - V_5[V_2\{1.1360\cos(\delta_5 - \delta_2) - 4.7725\sin(\delta_5 - \delta_2)\} \\ + V_7\{2.9540\cos(\delta_5 - \delta_7) - 7.4493\sin(\delta_5 - \delta_7)\}] = 0 \end{array} \tag{11}$$

$$\begin{array}{l} P_{G8} - 0.3 - V_8[V_6\{6.2893\cos(\delta_8 - \delta_6) - 22.0126\sin(\delta_8 - \delta_6)\} \\ + V_{28}\{1.4308\cos(\delta_8 - \delta_{28}) - 0.4499\sin(\delta_8 - \delta_{28})\}] = 0 \end{array} \tag{12}$$

$$P_{G11} - V_{11}[V_9\{-4.8077\sin(\delta_{11} - \delta_9)\}] = 0 \tag{13}$$

$$P_{G13} - V_{13}[V_{12}\{-7.1429\sin(\delta_{13} - \delta_{12})\}] = 0 \tag{14}$$

$$Q_{G1} - V_1[V_2\{5.2246\sin(\delta_1 - \delta_2) - 15.6467\cos(\delta_1 - \delta_2)\} + V_3\{1.2437\sin(\delta_1 - \delta_3) - 5.0960\cos(\delta_1 - \delta_3)\}] = 0 \tag{15}$$

$$Q_{G2} - 0.217 - V_2[V_1\{5.2246\sin(\delta_2-\delta_1) - 15.6467\cos(\delta_2-\delta_1)\} + V_4\{1.7055\sin(\delta_2-\delta_4) - 5.1974\cos(\delta_2-\delta_4) + V_5\{1.1360\sin(\delta_2-\delta_5) - 4.7725\cos(\delta_2-\delta_5)\} + V_6\{1.6861\sin(\delta_2-\delta_6) - 5.1165\cos(\delta_2-\delta_6)\}] = 0 \tag{16}$$

$$Q_{G5} - 0.942 - V_5[V_2\{1.1360\sin(\delta_5-\delta_2) - 4.7725\cos(\delta_5-\delta_2)\} + V_7\{2.9540\sin(\delta_5-\delta_7) - 7.4493\cos(\delta_5-\delta_7)] = 0 \tag{17}$$

$$Q_{G8} - 0.3 - V_8[V_6\{6.2893\sin(\delta_8-\delta_6) - 22.0126\cos(\delta_8-\delta_6)\} + V_{28}\{1.4308\sin(\delta_8-\delta_{28}) - 0.4499\cos(\delta_8-\delta_{28})] = 0 \tag{18}$$

$$Q_{G11} - V_{11}[V_9\{-4.8077\cos(\delta_{11}-\delta_9)\}] = 0 \tag{19}$$

$$Q_{G13} - V_{13}[V_{12}\{-7.1429\cos(\delta_{13}-\delta_{12})\}] = 0 \tag{20}$$

$$\begin{aligned} &50 \le P_{G_1} \le 200,\ 20 \le P_{G_2} \le 80,\ 15 \le P_{G_5} \le 50, \quad (\textit{Equality constraints}) \\ &10 \le P_{G_8} \le 35,\ 10 \le P_{G_{11}} \le 30,\ 12 \le P_{G_{13}} \le 40 \\ &-20 \le Q_{G_2} \le 100,\ -15 \le Q_{G_5} \le 80,\ -15 \le Q_{G_8} \le 60, \\ &-10 \le Q_{G_{11}} \le 50,\ -15 \le Q_{G_{13}} \le 60 \\ &0.95 \le V_{G_i} \le 1.1,\ i = 1,\ 2, 5, 8, 11, 13 \quad (\textit{Generator constraints}) \end{aligned} \tag{21}$$

$$0.85 \le V_{Li} \le 1.05,\ i = 2,\ 3, 4, 5, 7, 8, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 23, 24, 26, 29, 30 \quad (\textit{Load} - \textit{bus voltage constraint}). \tag{22}$$

- **The model under Run-2**

Simulation of overload with reduction of capacity of line 2 and 3 from 130 MW to 50 MW and 65 MW to 15 MW, respectively, is constructed as follows.

Find (S_2, S_3, P_{gi}, $i = 1, 2, 5, 8, 11, 13$) so as to

$$\textit{Minimize } Z = \left[P_1\left(\frac{1}{25-10}\right)d_1^-,\ P_2\left(\frac{1}{551942-530000}\right)d_2^-,\ P_3\left(\frac{1}{6.50-4.50}\right)d_3^-\right]$$

and satisfy

$$\mu_{F_1}: \frac{25 - \{(S_2-50)^2 + (S_3-15)^2\}}{25-10} + d_1^- - d_1^+ = 1,$$

$$\mu_{F_2}: \frac{551942 - (1650P_1 + 1.25P_1^2 + 1875P_2 + 1.425P_2^2 + 1875P_5 + 1.425P_5^2 + 2025P_8 + 1.5P_8^2 + 2025P_{11} + 1.5P_{11}^2 + 2025P_{13} + 1.5P_{13}^2}{551942 - 530000} + d_2^- - d_2^+ = 1,$$

$$\mu_{F_3}: \frac{6.50 - P_L}{6.50 - 4.00} + d_3^- - d_3^+ = 1,$$

subject to the system constraints in (9)–(22).

- **The model under Run-3**

Simulation of overload with outage of unit 3 at bus 5 with reduction of capacity of line 5 from 130 to 50 MW

Find $(S_5, P_{gi}, i = 1, 2, 5, 8, 11, 13)$ so as to

$$Minimize\, Z = \left[P_1\left(\frac{1}{5-2}\right)d_1^-,\ P_2\left(\frac{1}{571942 - 550000}\right)d_2^-,\ P_3\left(\frac{1}{10.00 - 8.50}\right)d_3^-\right]$$

and satisfy

$$\mu_{F_1}: \frac{5 - (S_{2-5} - 50)^2}{5 - 2} + d_1^- - d_1^+ = 1,$$

$$\mu_{F_2}: \frac{551942 - (1650P_1 + 1.25P_1^2 + 1875P_2 + 1.425P_2^2 + 2025P_8 + 1.5P_8^2 + 2025P_{11} + 1.5P_{11}^2 + 2025P_{13} + 1.5P_{13}^2}{551942 - 530000} + d_2^- - d_2^+ = 1,$$

$$\mu_{F_3}: \frac{10.00 - F_3}{10.00 - 8.50} + d_3^- - d_3^+ = 1,$$

subject to the given system constraints in (15)–(28).

The solutions achieved under three different runs are presented in Table 2.

Table 2 Simulation results under three runs

Run	Overloaded condition			Solution			
	Line/unit	MVA flow	MVA capacity	MVA flow	Power-loss (MW)	P_{ng} (ng = 1, 2, 5, 8, 11, 13)	Cost (Rs/hr)
1	Line 1	61.25	50	49.55	3.78	(82.27, 58.59, 50.00, 35.00, 30.00, 31.32)	536633.75
2	Line 2 and Line 3	35.44 20.30	50 15	35.50 14.95	4.50	(112.44, 20.46, 50.00, 35.00, 30.00, 40.00)	538219.96
3	Line 5 and Unit 3 out	40.89	50	**48.12**	8.76	(122.22, 31.24, 0.00, 35.00, 30.00, 40.00) (31.2 MW Load-Shaded)	550064.38

The results indicate that the decision achieved for each of the setup cases is a satisfactory one from the view point of distributing the proper power distribution in the decision-making environment.

5 Performance Analysis

To highlight the effective use of the proposed approach, the model solution is compared with the solution obtained using the particle swarm optimization (PSO). The resultant decision is presented in Table 3.

The diagrammatic representations of MVA flow and cost of CM under the two approaches are given in Figs. 1 and 2, respectively.

It is clear from the above that a better utilization of MVA capacity and management of congestion with comparatively less cost is achieved here in comparison with previously studied PSO-based approach.

Table 3 Simulation results under the PSO method

Case	Overloaded condition		Solution	
	Line/unit	MVA capacity	MVA flow	Cost (Rs/hr)
1	Line 1	50	49.16	541171
2	Line 2 and Line 3	50 15	12.31 14.99	542465
3	Line 5 and Unit 3 Out	50	49.88	565979

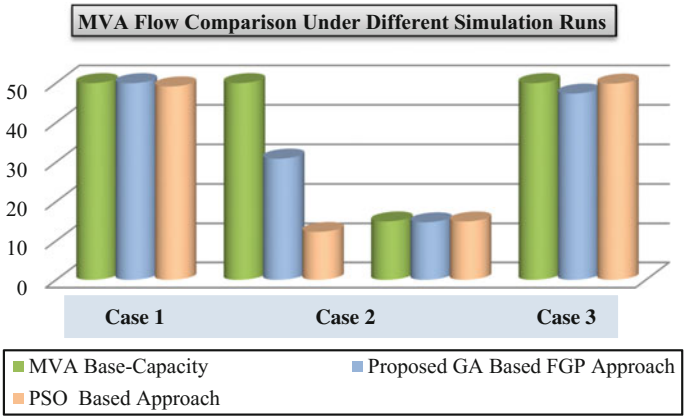


Fig. 1 MVA flow comparison

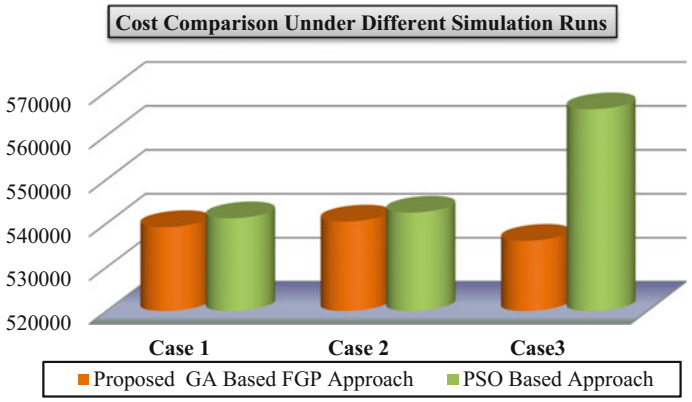


Fig. 2 Cost of congestion management comparison

6 Conclusions

A priority-based FGP approach for CM with multiple criteria is presented in this article. The main advantage of the proposed method is that optimization of various conflicting objectives according to their priorities of importance can be made for a unified management of congestions due to voltage instability and line overload in transmission network. Again, since GA is a population-based random search method, the demerit of using traditional single-point-based method for possible achievement of suboptimal solution can be avoided here in the decision environment. In future study, multilevel programming (MLP) approach (Sakawa and Nishizaki, 2009) [27] to CM problems in a hierarchical decision structure for sequential optimization of objectives on the basis of needs of a power plant operational system for OPF can be considered in a decision-making horizon.

References

1. Bhattacharya, K., Bollen, M., Daalder, J.E.: Operation of Restructured Power Systems. Springer, US (2001)
2. El Abiad, A.H., Stagg, G.S.: Automatic evaluation of power system performance-effect of line and transformer outages. AIEE Trans. **PAS-81**, 712–716 (1963)
3. Mamandur, K.R.C., Berg, G.J.: Economic shift in electric power generation with line flow constraints. IEEE Trans. Power Appar. Syst. **PAS-97**(7), 1618–1626 (1978)
4. Medicherla, T.K.P., Billinton, R., Sachdev, M.S.: Generation rescheduling and load shedding to alleviate line overload—analysis. IEEE Trans. Power Appar. Syst. **PAS-98**(6), pp. 1876–1884 (1979)
5. Finney, J.D., Othman, H.A., Rutz, W.L.: Evaluating transmission congestion constraints in system planning. IEEE Trans. Power Syst. **12**(3), 1143–1149 (1997)

6. Fang, R.S., David, A.K.: An integrated congestion management strategy for real time system operation. IEEE Power Eng. Rev. **19**(5), 52–53 (1999)
7. Wang, X., Song, Y.H.: Advanced real-time congestion management through both pool balancing market and bilateral market. IEEE Power Eng. Rev. **20**(2), 47–49 (2000)
8. Rau, N.S.: Transmission loss and congestion cost allocation: an approach based on responsibility. IEEE Trans. Power Syst. **15**(4), 1401–1409 (2000)
9. Fuerte-Esquivel, C.R., Acha, E., Ambriz-Perez, H.A.: Thyristor controlled series compensator model for the power flow solution of practical power networks. IEEE Trans. Power Syst. **15** (1), 58–64 (2000)
10. Lee, K.-H.: Optimal siting of TCSC for reducing congestion cost by using shadow prices. Int. J. Electr. Power Energy Syst. **24**, 647–653 (2002)
11. Rodrigues, A.B., DaSilva, M.G.: Impact of multilateral congestion management on the reliability of power transactions. Int. J. Electr. Power Energy Syst. **25**, 113–132 (2003)
12. Milano, F., Canizares, C.A., Invernizzi, M.: Multi-objective optimization for pricing system security in electricity markets. IEEE Trans. Power Syst. **18**(2), 596–604 (2003)
13. Kumar, A., Srivastava, S.C., Singh, S.N.: A zonal congestion management approach using real and reactive power rescheduling. IEEE Trans. Power Syst. **18**(1), 554–562 (2004)
14. Talukdar, B.K., Sinha, A.K., Mukhopadhyay, S., Bose, A.: A computationally simple method for cost-efficient generation rescheduling and load shedding for congestion management. Int. J. Electr. Power Energy Syst. **27**(5–6), 379–388 (2005)
15. Pearl, J.: HeuristICS: Intelligent Search Strategies for Computer Problem Solving. Addison-Wesley, New York (1983)
16. Hazra, J., Sinha, K.A.: Congestion management using multi objective particle swarm optimization. IEEE Trans. Power Syst. **22**(4), 1726–1734 (2007)
17. Dutta, S., Singh, S.P.: Optimal rescheduling of generator for congestion management based on particle swarm intelligence. IEEE Trans. Power Syst. **23**(4), 1560–1569 (2008)
18. Balaraman, S., Kamaraj, N.: Congestion management using hybrid particle swarm optimization technique. Int. J. Swarm Intell. Res. **1**(3), 51–66 (2010)
19. Hazra, J., Sinha, A.K.: Identification of catastrophic failures in power system using pattern recognition and fuzzy estimation. IEEE Trans. Power Syst. **24**(1), 378–387 (2009)
20. Nouri, H., Tang, S.H., Tuah, B.T.H., Anuar, M.K.: BASE: A bacteria foraging algorithm for cell formation with sequence data. J. Manuf. Syst. **2–3**, pp. 102–110 (2010)
21. Venkaiah, C.H., Vinod Kumar, D.M.: Fuzzy adaptive bacterial foraging congestion management using sensitivity based optimal active power re-scheduling of generator. Appl. Soft Comput. **11**(8), 4921–4930 (2011)
22. Vijaya Kumar, J., Vinod Kumar, D.M., Edukondalu, K.: Strategic bidding using fuzzy adaptive gravitational search algorithm in a pool based electricity market. Appl. Soft Comput. **13**(5), 2445–2455 (2013)
23. Awerbach, S., Ecker, J.G., Wallace, W.A.: A note: hidden nonlinearities in the application of goal programming. Manage. Sci. **22**, 918–920 (1976)
24. Pal, B.B., Sen, S., Kumar, M.: A linear approximation approach to chance constrained multiobjective decision making problems. In: IEEE Xplore, pp. 70–75 (2009). Online: IEEEDigital Library, doi:10.1109/ICADVC.2009.5378202
25. Goldberg, D.E.: Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley Reading, MA (1989)
26. Zimmermann, H.-J.: Fuzzy Sets, Decision Making and Expert Systems. Kluwer Academic Publisher, Boston (1987)
27. Sakawa, M., Nishizaki, I.: Cooperative and Non-cooperative Multi-level Programming. Springer, New York (2009)

Effect of Surface-to-Volume Ratio on Eigenenergy in Quantum Ring

Swapan Bhattacharyya, Sourish Haldar and Arpan Deyasi

Abstract Eigenenergies of lowest three quantum states in semiconductor quantum ring is analytically determined as a function of surface-to-volume ratio subject to the application of external electric field applied along the axis of the ring, which is perpendicular to ring plane. n-GaAs material is considered for simulation, and the results are compared with that obtained in the absence of field. Result shows that with increasing S/V ratio, energy increases almost linearly in the presence of field, whereas the behavior is nonlinear when field is absent. Critical finding in this respect is that intersubband transition remains constant with increasing S/V ratio, which speaks about its candidature as optical emitter/detector within dimensional variation over nanorange. Variation is also calculated for different surface areas and volume of nanodevice. Result speaks about possible tuning of the eigenenergy by external field in IR wavelength region.

Keywords Surface-to-volume ratio · Eigenenergy · Transition energy · Electric field · Structural parameters

S. Bhattacharyya (✉)
Department of Electronics and Communication Engineering,
JIS College of Engineering, Kalyani, India
e-mail: swapanbhattacharyya@ieee.org

S. Haldar
Department of Electronics and Communication Engineering,
NIT Durgapur, Durgapur, India
e-mail: sourish@gmail.com

A. Deyasi
Department of Electronics and Communication Engineering,
RCC Institute of Information Technology, Kolkata, India
e-mail: deyasi_arpan@yahoo.co.in

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_8

1 Introduction

Theoretical [1, 2] and experimental [3, 4] investigations of semiconductor nanostructures precisely, quantum dots (ring, box, dash, disk) and quantum wires are focused nowadays due to the faster technological advancement in the fabrication techniques [5]. In nanostructure device, one or more dimensions are comparable to or less than electron wavelength (de-Bröglie wavelength which is few nm) and as a consequence of that, electron motion along the quantum confined dimensions becomes restricted. This confinement in quantum scale leads to quantization of the energy levels (named as eigenenergy) and makes meaningful changes in microscopic and mesoscopic properties of the structure [6]. Semiconductor nanostructures with ring [7] and disk [8] shaped geometry have received a considerable attention due to their electronic and optical properties. A quantum dot structure with circular geometry with a 'hole' at the middle is termed as quantum ring, and these particular structures are under intense study [9, 10] due to the possibility of experimental observations of the Aharonov–Bohm effect [11]. Obviously, eigenstates of ring structure becomes a function of their dimension, and can very easily be tuned by external electric field, applied along the axis of the dot [12, 13]. This characteristic makes them a promising candidate for optical emitter/detector in IR region. In this context, surface-to-volume ratio of the device can play a major role in determining the optical performance.

For nanostructures, surface area-to-volume ratio greatly influences electrical and optical properties. For a given structure, when the total volume is divided into smaller number of elements, total surface area increases. Henceforth, with reduction of particle dimension, more number of atoms are found at the surface compared to the atoms inside the structure. Henceforth, nanodevices have much greater surface area per unit volume when compared with their bulk counterparts, exhibits enhanced chemical activity, which leads to the way of making more promising devices for significant applications [14].

In this paper, a theoretical model is presented for investigating lowest three eigenstates and corresponding intersubband transition energy as a function of surface area, volume of the device and surface-to-volume ratio. Electric field is applied to tune the subband states. Result plays significant role in determining performance of the device as optical emitter/detector.

2 Mathematical Modeling

We consider the ring-shaped geometry as shown in Fig. 1.

Along the axis (perpendicular plane) of quantum ring, external electric field (F) is applied. To solve the problem, we consider cylindrical co-ordinate system (ρ, θ, z) [15]. Therefore, the time-independent Schrödinger equation with modified potential function for the wavefunction Ψ may be put as

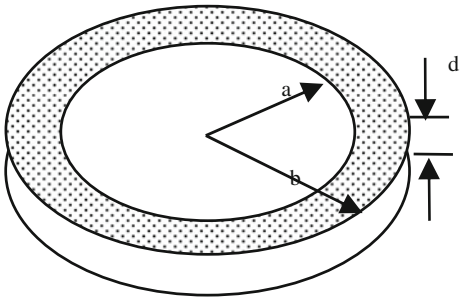


Fig. 1 Schematic representation of a quantum ring

$$-\frac{\hbar^2}{2m^*}\left[\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\Psi}{\partial\rho}\right)+\frac{1}{\rho^2}\frac{\partial^2\Psi}{\partial\rho^2}+\frac{\partial^2\Psi}{\partial z^2}\right]-eF\Psi=E\Psi \tag{1}$$

with the boundary conditions a $\leq$ ρ $\leq$ b and z $\leq$ d.

LHS of Eq. (1) may be considered as the effective Hamiltonian operator ($\hat{H}$) in cylindrical coordinate system working on the wavefunction (Ψ) [16]. For solution of Eq. (1), separation of variable method can be adopted.

In order to calculate energy states for the structure, Eq. (2) needs to be solved which is in determinant form

$$\begin{vmatrix} J_{1/3}\left(\frac{2}{3}\varsigma_1^{3/2}\right) & J_{-1/3}\left(\frac{2}{3}\varsigma_1^{3/2}\right) \\ J_{1/3}\left(\frac{2}{3}\varsigma_2^{3/2}\right) & J_{-1/3}\left(\frac{2}{3}\varsigma_2^{3/2}\right) \end{vmatrix}=0 \tag{2}$$

where ζ_1 (ζ_2) corresponds to the value of ζ given by

$$\varsigma=\left(\frac{2m^*eF}{\hbar^2}\right)^{1/3}\left(z+\frac{E_{zn}}{eF}\right) \tag{3}$$

calculated at the inner boundary (intersection of air and material) and outer boundary (intersection of material and air) boundaries where $J_{-1/3}\,(J_{1/3})$ is the Bessel functions of order (−1/3) and (1/3).

Equation (4)

$$\begin{vmatrix} J_m(\lambda_1) & Y_m(\lambda_1) \\ J_m(\lambda_2) & Y_m(\lambda_2) \end{vmatrix}=0 \tag{4}$$

can be solved to obtain radius vector-dependent component of the energy eigenvalue ($\hbar^2\lambda^2/2m^*$), where λ_1, λ_2 correspond to the value of λ at two boundaries [16].

The solution of (4) requires two quantum numbers, *m* and *l* (say), which indicates that for *m*-th order of Bessel function, *l*-th zero is the solution. The solution of Eq. (2), however, gives rise to a single quantum number (*n*, say). Thus, the energy eigenvalue is quantized with three quantum numbers.

Using Eqs. (3) and (4), energyvalues for the confined states of the quantum ring can be evaluated and approximately written as

$$E_{lmn} = \frac{\hbar^2}{2m^*}\lambda_{ml}^2 + \sqrt{\frac{\hbar^2}{2m^*}\left(\frac{n\pi}{d}\right)^2 \cdot \left\{\frac{\hbar^2}{2m^*}\left(\frac{n\pi}{d}\right)^2 - eFd\right\}} \tag{5}$$

for not very high electric field. Here we consider the fact that phase part of electric field propagates through uniform medium. Henceforth, propagation matrix may be written as the function of well and barrier widths.

3 Results and Discussions

Using Eq. (5), energy for lowest thee eigenstates are computed and plotted as a function of surface area and surface-to-volume ratio. Figure 2 shows the energy states in absence and presence of electric field. From the plot, it is seen that energy monotonically decreases with increasing surface area. But the rate of reduction is higher for lower surface area and reduces when area becomes large. This is due to the fact that with increasing surface area, confinement becomes weaker, and thus energy becomes almost constant. It may also be noted that application of electric field perpendicular to the structure reduces the energy value, and the reduction is higher for higher surface area.

In Fig. 3, energies of eigenstates are plotted with surface-to-volume ratio. From the plot, it is observed that energy monotonically increases with increment of the ratio. However, the change is linear in presence of electric field, whereas it is slightly nonlinear when field is absent. Nonlinearity increases for higher value of the ratio. Hence it is evident that for lower S/V ratio, energy is far lower than that

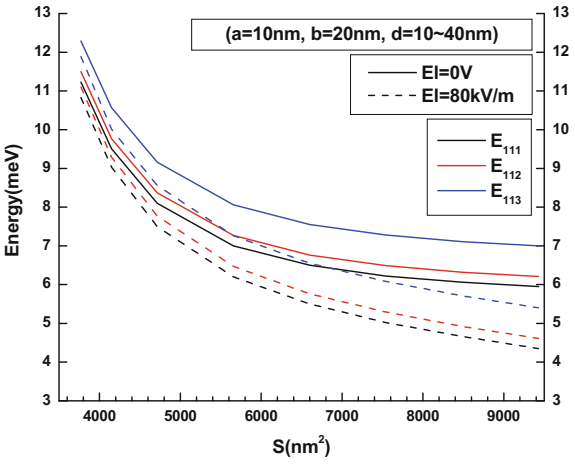


Fig. 2 Energy for lowest three states with surface area

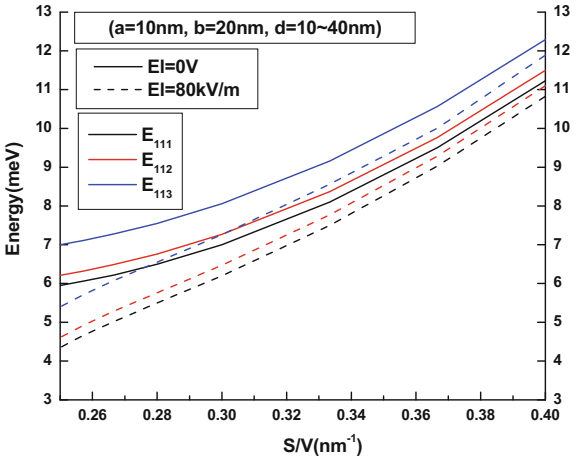


Fig. 3 Energy for lowest three states with increasing surface-to-volume ratio

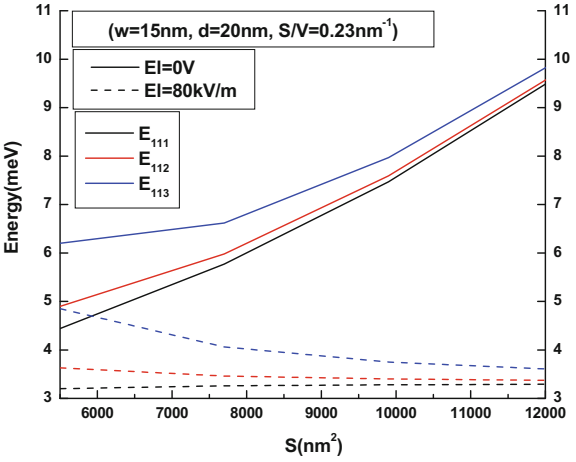


Fig. 4 Energy for lowest three states with increasing surface area for constant surface-to-volume ratio

obtained for the structure when it is not under external excitation. Also for higher S/V ratio, the difference between the energy value becomes less considering the effect of electric field.

Figure 4 shows the variation of quantum confined states with increasing surface area for a constant surface-to-volume ratio. It is interesting to note that with increasing area for constant S/V factor, energy monotonically increases in absence of filed, but reduces and becomes almost constant when field is present. This result has importance in nanodevice design. The result signifies the fact that for constant S/V ratio, if surface area is increased, tuning of energy becomes more prominent. Thus for larger change in energy, higher surface area is preferred, whereas for lower tuning range, lower cross section is important. This drastic change may play significant role in determining its candidature for future application.

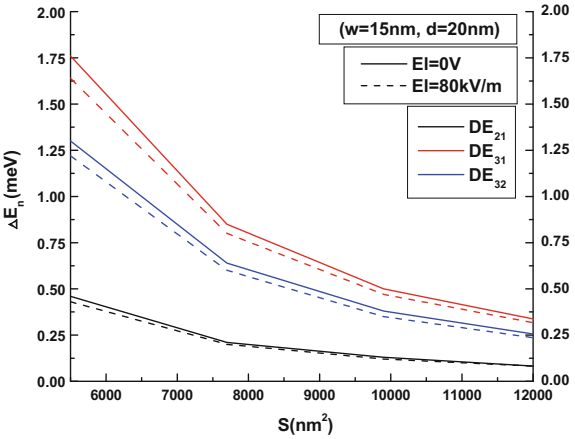


Fig. 5 Intersubband transition energy for lowest three states with surface area

Table 1 Transition energy with constant S/V ratio in the presence and absence of efield

S/V	Transition energy (eV)					
	Field = 0			Field = 80 kV/m		
	ΔE_{21}	ΔE_{31}	ΔE_{32}	ΔE_{21}	ΔE_{31}	ΔE_{32}
0.233	0.46	1.76	1.30	0.43	1.64	1.22
0.233	0.21	0.85	0.64	0.20	0.80	0.60
0.233	0.13	0.50	0.38	0.12	0.47	0.35
0.233	0.08	0.33	0.25	0.08	0.31	0.23

Intersubband transition energy for the lowest three quantum states is calculated as a function of surface area in Fig. 5. From the plot, it is observed that higher transition energy is obtained for lower surface area, and it monotonically reduces with increasing the area. However, the change is significant when transition is considered between non-adjacent states, more precisely, between second excitation state and ground state. Also the rate of reduction of transition energy is significant for lower surface area. It may also be noted in this context that transition energy is lower when field is applied.

For constant S/V ratio, we have calculated the variation of transition energy. Result is noted in Table 1. It is observed that with increasing surface area, transition energy decreases. This is observed both in absence and presence of field. However, the change is higher when field is absent. It may be mentioned in this context that though eigenenergy increases with increasing surface area for constant S/V ratio in absence of field, but transition energy decreases. This is quite interesting form application standpoint. The rate of decrement is almost quite same for both the cases.

4 Conclusion

Quantized carrier states of electron for different surface areas and surface-to-volume ratio of a quantum ring of moderate band gap semiconductor have been analytically calculated and result is graphically represented in the presence of electric field. Result shows that eigenenergy decreases with increasing surface area, but increases S/V ratio. For constant S/V ratio, energy enhances in the absence of field, whereas decreases in the presence of field. This shows the importance of the present analysis. For optical emitter/detector application, tuning of energy is very much essential by external excitation, and by virtue of electric field, energy can be tuned keeping a note to the S/V ratio. Thus overall dimension of the device plays crucial role in determining its candidature for optical applications.

References

1. Urban, D., Braun, M., König, J.: Theory of a magnetically controlled quantum-dot spin transistor. Phys. Rev. B **76**, 125306 (2007)
2. Zhang, L., He, J., Zhang, J., Liu, F., Fu, Y., Song, Y., Zhang, X.: An analytic model for nanowire MOSFETs with Ge/Si core/shell structure. IEEE Trans. Electron Devices **55**, 2907–2917 (2008)
3. Song, K.K., Lee, S.: Highly luminescent (ZnSe) ZnS core-shell quantum dots for blue to UV emission: synthesis and characterization. Curr. Appl. Phys. **1**, 169–173 (2001)
4. HuEmail, M.Z., Zhu, T.: Semiconductor nanocrystal quantum dot synthesis approaches towards large-scale industrial production for energy applications. Nanoscale Res. Lett. **10**, 469 (2015)
5. Prasad, M., Verma, P., Deyasi, A.: Computing surface potential of submicron MOSFET with thin dielectric. In: National Conference on Frontline Research in Computer, Communication and Device, 149 (2015)
6. Gangopadhyay, S., Nag, B.R.: Energy levels in three-dimensional quantum-confinement structures. Nanotechnology **8**, 14–17 (1997)
7. Huang, G., Guo, W., Bhattacharya, P., Ariyawansa, G., Perera, A.G.U.: A quantum ring terahertz detector with resonant tunnel barriers. Appl. Phys. Lett. **94**, 101115 (2009)
8. Zhou, W., Pennycook, S.J., Idrobo, J.C.: Probing the electronic structure and optical response of a graphene quantum disk supported on monolayer grapheme. J. Phys. Condens. Matter **8** (24), 314213 (2012)
9. Deyasi, A., Bhattacharyya, S., Das, N. R.: Effect of shell thickness on intersubband transition energies in GaAs/$Al_xGa_{1-x}As$ Inverted Core-Shell Nanodot. In: Springer Proceedings in Physics: Advanced Nanomaterials and Nanotechnology, vol. 143, no. 55, p. 551 (2013)
10. Deyasi, A., Bhattacharyya, S.: Effect of band nonparabolicty on electron energies of a quantum disk in presence of electric field. INDICON **13**, 1 (2013)
11. Chena, G.Y., Chenb, Y.N., Chuuc, D.S.: The Aharonov-Bohm effect in concentric quantum double rings. Solid State Commun. **143**, 515–518 (2007)
12. Acosta, R.E., Morales, A.L., Duque, C.M., Mora-Ramos, M.E., Duque, C.A.: Optical absorption and refractive index changes in a semiconductor quantum ring: electric field and donor impurity effects. Phys. Status Solidi (b) **253**, 744–754 (2016)
13. Bhattacharyya, S., Deyasi, A., Das, N.R.: Effect of band nonparabolicity and material composition on intersubband transition energies of a heterostructure quantum ring for optical transmitter. In: National Conference on Computer and Communication Systems, pp. 1–5, Nov 2012

14. Kundu, P., Ghosh, P., Deyasi, A.: Analytical computation of absorption coefficient for intersubband transition in MQW structure. In: Lecture Notes in Electrical Engineering: Computational Advancement in Communication Circuits and Systems, part 6: Advances in Devices and Circuit, vol. 335, no. 35, p. 321 (2014)
15. Bhattacharyya, S., Deyasi, A., Das, N.R.: Effect of material composition on eigenstates of a three-dimensional heterostructure quantum ring in presence of electric field. INDICON **2012**, 649–652 (2012)
16. Bhattacharya, S., Sen, S., Das, N.R.: Tuning of intersubband transition in a quantum ring by external; electric field. In: International Workshop on the Physics of Semiconductor Devices, IWPSD (2007)

Part II
Nanotechnology and Applications

Unusual Conductance Characteristics in Single Electron Transistor

Arkadeep Paul, Ritabrata Chakraborty, Arpan Deyasi and Shrabani Nayak

Abstract Dependence of conductance on equivalent circuit parameters in single electron transistor is analytically computed for electrical performance estimation. Distorted conductance profiles are obtained when a few passive components exceeds threshold limit, and negative spikes are also possible, as revealed from simulation. Steady-state master equation is solved with appropriate boundary conditions when source and drains are connected via quantum dot, which ensures tunneling process. Fermi Golden Rule is applied to calculate probabilistic values of all stochastic processes and effect of source and drain resistances and capacitances as well as gate capacitance are considered for determining conductance. Simulated findings are important for practical application of SET as infrared detector and charge sensor.

Keywords Single electron transistor · Conductance · Coulomb blockade · Quantum dot · Distorted profile · Negative spikes

1 Introduction

The continuous shrinkage in device size, precisely the channel length in MOSFET leads to various short channel effect, which cannot be counterbalanced after a threshold limit [1]. This leads to emerge new devices with extremely low power

A. Paul (✉) · R. Chakraborty · A. Deyasi · S. Nayak
Department of Electronics and Communication Engineering,
RCC Institute of Information Technology, Kolkata, India
e-mail: arkadeep.paul@gmail.com

R. Chakraborty
e-mail: ritobroto.bms@gmail.com

A. Deyasi
e-mail: deyasi_arpan@yahoo.co.in

S. Nayak
e-mail: shrabani.communication@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_9

dissipation feature and controllable current–voltage characteristics; and single electron transistor becomes a prime candidate as it satisfies the requirements for memory [2], detector [3] and sensor [4] applications. It is a transistor where source and drain junctions are quantum mechanically connected via quantum dot; and current is flown via tunneling mechanism overcoming Coulomb Blockade.

For measuring current flow in SET, estimation of charge in quantum dot under biasing condition is required, which is experimentally demonstrated by Berman and others [5]. It can be used for multiple-valued logic system [6]. Uchida first proposed the compact model for SET [7] assuming symmetric device and excluding background effect. Later, these effects are included by Inokawa [8]. But both the works are limited by the condition $V_{DS} < q/C$. This restriction is removed by master equation-based technique as individually proposed by Royer [9], Lientschnig [10]. Later, Mahapatra et al. [11] showed the improvement of the master equation technique which becomes applicable for both analog and digital circuits. Shahhosei [12] also carried out work by considering metallic island SET, after which kink effect [13] is incorporated. Bohm [14] developed adaptive Monte Carlo algorithm for simulation of single electron transistor, which provides both DC characteristics as well as tunelling dynamics.

In this paper, conductance characteristics are analytically computed considering the tunneling mechanism as stochastic process and solving the rate equations using Fermi Golden Rule. Effect of source and drain resistances and capacitances, along with gate capacitance is analyzed on conductance, and unusual behavior is observed when a few circuit parameters exceed threshold value. Negative spikes and distorted profiles are obtained which speaks about the limit of passive components for circuit operation.

2 Mathematical Formulation

Tunneling of charge 'q' from source to quantum dot or from dot to drain is a probabilistic process, and thus is described by probability distribution function as

$$\frac{\partial \rho(N,t)}{\partial t} = \rho(N+1)[\Gamma_S^+(N+1) + \Gamma_D^-(N+1)] - \rho(N)[\Gamma_S^-(N) + \Gamma_D^+(N)], \quad (1)$$

where Γ_S and Γ_D are source and drain tunneling rates respectively. These rates are computed using Fermi's Golden Rule as

$$\Gamma_D^{\pm}(N) = \frac{1}{R_D q^2}\left[\frac{-\Delta F_D^{\pm}}{1 - \exp(\Delta F_D^{\pm}/k_B T)}\right] \quad (2.1)$$

$$\Gamma_S^{\pm}(N) = \frac{1}{R_S q^2}\left[\frac{-\Delta F_S^{\pm}}{1 - \exp(\Delta F_S^{\pm}/k_B T)}\right] \quad (2.2)$$

Here ΔF denotes change of free energy, and is calculated when electron tunnels across tunnel capacitances. It depends on the number of excess electrons inside quantum dot, and are given by

$$\Delta F_D^{\pm}(n_D, n_S) = \frac{q}{C_\Sigma}\{\frac{q}{2} \pm (N_e - Q_0) \mp (C_G + C_S)V \pm C_G V_G\} \tag{3.1}$$

$$\Delta F_S^{\pm}(n_D, n_S) = \frac{q}{C_\Sigma}\{\frac{q}{2} \mp (N_e - Q_0) \mp C_D V \mp C_G V_G\} \tag{3.2}$$

Equation (1) is solved by setting probability density as zero, so that it can be put into the form

$$\rho(N)[\Gamma_S^-(N) + \Gamma_D^+(N)] = \rho(N+1)[\Gamma_S^+(N+1) + \Gamma_D^-(N+1)] \tag{4}$$

Equation (4) can be solved by introducing boundary condition as

$$\rho(N) \to 0, N \to \pm\infty \tag{5.1}$$

The distribution function has also normalization condition

$$\sum_{N=-\infty}^{\infty} \rho(N) = 1. \tag{5.2}$$

Applying suitable boundary condition, final expressions for current are given by

$$I(V) = q \sum_{N=-\infty}^{\infty} \rho(N)[\Gamma_D^+(N) - \Gamma_D^-(N)] \tag{6.1}$$

$$I(V) = q \sum_{N=-\infty}^{\infty} \rho(N)[\Gamma_S^+(N) - \Gamma_S^-(N)], \tag{6.2}$$

where Eqs. (6.1) and (6.2) represents current at drain and source ends respectively.

3 Results and Discussions

Figure 1 shows the conductance profile of single electron transistor, where peaks are decreasing with increase of forward or reverse bias. In this case, equivalent circuit parameters are chosen in such a way that near ideal profiles can be obtained.

With increase of C_D, conductance profile tends to become flatter with increase in the value of C_D. However, at the same time, the amplitude of the conductance profile increases as C_1 is increased. It may be noted that at sufficiently high values of C_D, keeping the other tunneling parameters constant, two symmetric downward

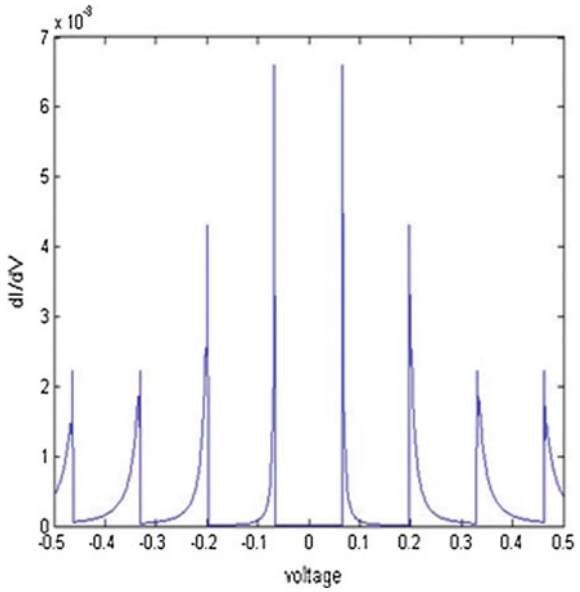


Fig. 1 Conductance profile with change of applied bias

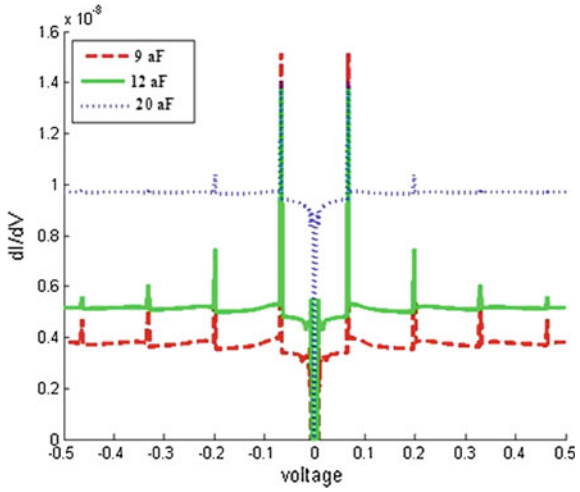


Fig. 2 Conductance profiles for three sufficiently high values of C_D

spikes appears on both sides of V = 0. With increase in the value of C_D, these two spikes come closer and converge at V = 0.

With increase in C_S, the number of spikes across the length of x-axis is found to increase drastically. As C_S increases, the spikes tend to be congested around V = 0 on the voltage axis, as can be seen from Fig. 2. The amplitude of the spikes increases with the increase in the value of C_S, although area of spikes decreases simultaneously. It has also been observed that for a given set of parametric values,

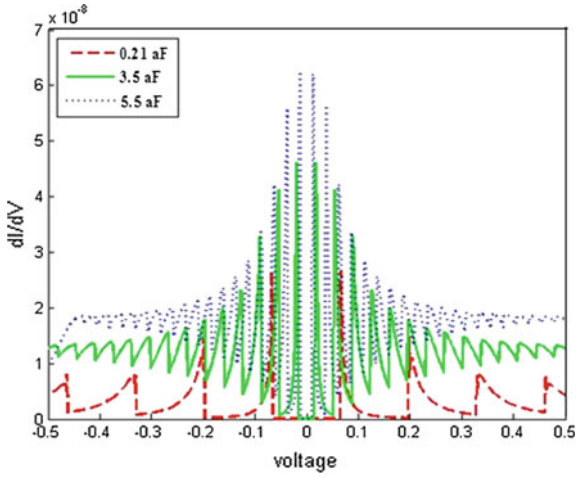


Fig. 3 Conductance profiles for three sufficiently high values of C_S

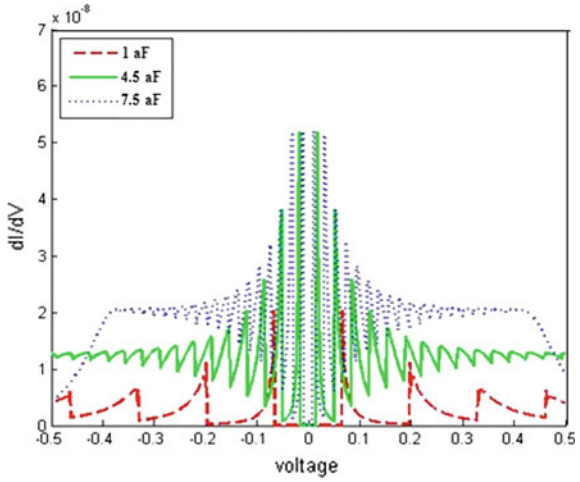


Fig. 4 Conductance profiles for three sufficiently high values of R_D

if we go on reducing C_S, keeping the other parameters constant, then a point will be reached, below which no further change in conductance profile is observed, even if C_S is reduced further. For our analysis, it has been found that at approximately 1 aF $< C_S <$ 2aF the conductance profile gets fixed and remains constant, for further reduction in value of C_S.

With increase in C_g, the number of spikes across the length of x-axis increases greatly. As C_g increases, the spikes tend to be congested around V = 0 on the voltage axis, as can be seen from Fig. 4. However, the variation of conductance profile, for C_g, is more rapid and faster than the variation for C_S, as can be observed by comparing Figs. 3 and 4. The amplitude of the spikes increases with the increase in the value of C_g, although area of spikes decreases with increasing C_g. But we

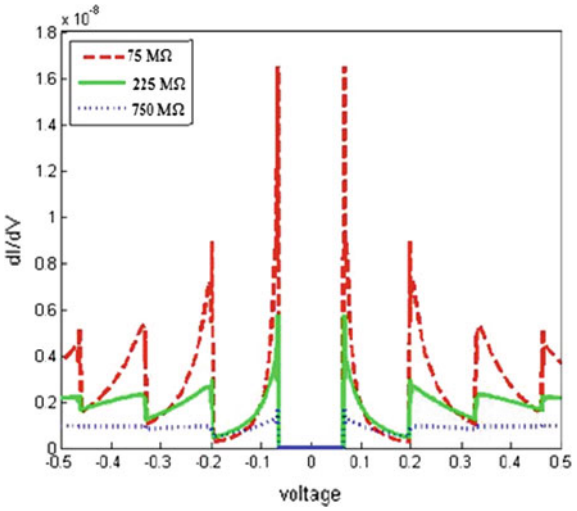


Fig. 5 Conductance profiles for three sufficiently high values of R_S

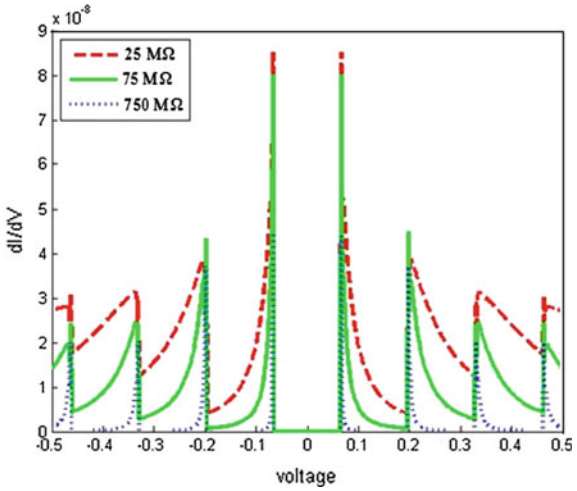


Fig. 6 Conductance profiles for three sufficiently high values of R_D

should also note that the amplitude of the spikes, which are nearer to V = 0, remain practically unaltered for varying values of C_g.

The amplitude of the conductance spikes reduces with increase in R_D, as can be seen from Fig. 5. At very high values of R_1, the entire conductance profile tends to become constant with voltage. With decrease in the value of R_D, the area of the spikes reduces and they tend to get completely discretized, as shown in Fig. 5. However, at very low values of R_D, the spikes for the positive values of voltage a paper to be shorter than those on the negative values of V. Such a profile is shown in Fig. 5.

The amplitude of the conductance spikes reduces with increase in R_D, as can be seen from Fig. 6. At very high values of R_D, the entire conductance profile tends to

become constant with voltage. With decrease in the value of R_D, the area of the spikes reduces and they tend to get completely discretized, as shown in Fig. 6. However, at very low values of R_D, the spikes for the positive values of voltage appear to be shorter than those on the negative values of V.

4 Conclusion

We have ventured to analyze the electrical characteristics of SET, as a function of the tunneling parameters. In this process, in addition to the I–V profile, we have also taken into consideration the conductance profile of SET and have performed the same operations on the conductance profile as well. It is noteworthy to mention that that SET being a nanodevice, it involves a complex process to analyze its properties by hardware setup. Hence we have devised an appropriate code, based on a standard set of equations, which establishes a relation between tunneling parameters and electrical characteristics.

References

1. ITRS Roadmap (2007)
2. Yano, K., Ishii, T., Hashimoto, T., Kobayashi, T., Murai, F., Seki, K.: Room-temperature single-electron memory. IEEE Trans. Electron Devices **41**, 1628–1638 (1994)
3. Astafiev, O., Antonov, V., Kutsuwa, T., Komiyama, S.: Single electron transistors as far-infrared photon detectors. In: Device Research Conference 145–147 (2001)
4. Knobel, R.G., Cleland, A.N.: Nanometre-scale displacement sensing using a single electron transistor. Nature **424**, 291–293 (2003)
5. Berman, D., Zhitenev, N.B., Ashoori, R.C., Smitha, H.I., Melloch, M.R.: Single-electron transistor as a charge sensor for semiconductor applications. J. Vac. Sci. Technol. B **15**, 2844–2847 (1997)
6. Degawa, K., Aoki, T., Higuchi, T., Inokawa, H., Takahashi, A.: A single-electron-transistor logic gate family and its application—part i: basic components for binary, multiple-valued and mixed-mode logic. In: 34th International Symposium on Multiple-Valued Logic, pp. 262–268 (2004)
7. Uchida, K.: Analytical single-electron transistor (SET) model for design and analysis of realistic SET circuits. Jpn. J. Appl. Phys. Part 1 **39**, 2321–2324 (2000)
8. Inokawa, H., Takahashi, Y.: A compact analytical model for asymmetric single-electron transistors. IEEE Trans. Electron Devices **50**(2), 455–461 (2003)
9. Le Royer, C.: Accurate modeling of quantum-dot based multi-tunnel junction memory. In: Proceeding of ESSDERC, pp. 403–406 (2002)
10. Lientschnig, G., Weymann, I., Hadley, P.: Simulating hybrid circuits of single-electron transistors and field-effect transistors. Jpn. J. Appl. Phys. Part 1 **42**, 6467–6472 (2003)
11. Mahapatra, S., Vaish, V., Wasshuber, C., Banerjee, K., Ionescu, A.M.: Analytical modeling of single electron transistor for hybrid CMOS-SET analog IC design. IEEE Trans. Electron Devices **51**, 1772–1782 (2004)

12. Shahhoseini, A., Saghafi, K., Moravvej-Farshi, M.K., Faez, R.: An analytic model for kink effect in i–v characteristics of single electron transistor. Iran. J. Electr. Electron. Eng. **5**, 234–243 (2009)
13. Hasaneen, E.-S.A.M., Wahab, M.A.A., Ahamed, M.G.: Exact analytical model of single electron transistor for practical ic design. Microelectron. Reliab. **51**, 733–745 (2011)
14. Bohm, D.: Quantum Theory, Prentice Hall, Jan (1989)

Conductivity, Dielectric and Impedance Property Analysis of $Y_{1-X}La_xCrO_3$ Nanoparticles

R. Sinha, S. Basu and A.K. Meikap

Abstract Perovskite-type La-doped Yttrium Chromite ($Y_{1-X}La_xCrO_3$) nanoparticles are prepared via sol–gel method. The formation of the compound is confirmed by X-ray diffraction (XRD) technique. Particle sizes are calculated by Transmission Electron Microscopy (TEM) measurements. Electrical DC conductivity is studied. DC conductivity analysis shows that all the samples have semiconducting behavior. Dc activation energy decreases when the lanthanum doping content is increased. The dielectric properties of all the samples are studied within the temperature range 298–523 K and in the frequency range 20–1 MHz. Impedance (AC) analysis suggests that the assistance of grain boundary governs over the grain.

Keywords Nanoparticles · Electrical conductivity · Dielectric properties · Impedance analysis

1 Introduction

In the recent times orthochromites having formula $RCrO_3$ (where R = Y or any rare earth element) [1] are widely investigated for showing large variation of physical and chemical features. These properties depend on the rare earth ion of the perovskites configuration. In the recent years many investigations were performed on understanding the magnetic and electrical properties of these materials. As they are chemically stable at higher temperature there is a growing interest of application in the field of high temperature electrode and thermoelectric materials and solid oxide fuel cell. Recently, this group of compounds has drawn much more concern for the combination of ferroelectric and ferromagnetic orderings in the same volume of the

R. Sinha (✉)
Department of Physics, Asansol Engineering College, Asansol 713305, West Bengal, India
e-mail: sinha.ranjita15@gmail.com

R. Sinha · S. Basu · A.K. Meikap
Department of Physics, NIT Durgapur, Durgapur 713209, West Bengal, India
e-mail: meikapnitd@yahoo.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_10

same material. The coupling between ferroelectric and ferromagnetic orderings in the same phase is much more important with respect to the view of basic science and also in the application field such as spintronic and storing data devices. Thermally, electrically and chemically $YCrO_3$ is very stable. It also has structural stability for its high melting point temperature [2]. Yttrium Chromite also has high electrical conductivity. Recently, many investigations are focused on $YCrO_3$ due to its biferroic [3] property that means uniting ferroelectricity and ferromagnetism in the same phase within same material. In the present work we are addressing to the electrical behavior of nanocrystalline $YCrO_3$ samples doped with La and here we have achieved higher electrical conductivity which is very much important for its application in solid oxide fuel cells.

2 Experimental Technique

Nanocryatalline La-doped $Y_{1-X}La_xCrO_3$ samples (x = 0, 0.01, 0.05, 0.10) are produced using sol–gel procedure. Certain quantity of Nitrate precursor of Yttrium [$Y(NO_3)_3$, $6H_2O$], Chromium [$Cr(NO_3)_3$, $9H_2O$] and Lanthanum [$La(NO_3)_3$, $6H_2O$] are dissolved in 50 ml distilled water. In the next step 20 ml PVA solution (strength 12.5 gm/l) is added to the previous solution. Then it is kept under continuous stirring for 1 h. At the end the solution is evaporated and the anhydrous powder is heated at 800 °C during 1 h.

3 Results and Discussions

3.1 XRD and TEM Analysis

The powder X-ray diffraction (XRD) measurements shows that all the samples are well crystalline in nature. Figure 1 depicts the XRD patterns for pure $YCrO_3$ sample and 10% La doped sample. The prominent peaks in Fig. 1 indicates various (*hkl*) planes and it is matching with "ICDD PDF No. 34-0365." The crystallite size (average) is estimated following the Debye Scherer equation [4, 5]. It can be written as,

$$D = 0.9\lambda/\beta\cos\theta \tag{1}$$

In Eq. 1 D, λ and β are the crystallite size, X-ray wavelength and FWHM for the diffraction peak respectively. Average crystallite size calculated is 40 nm. Rietveld analysis shows that $Y_{1-X}La_xCrO_3$ compound becomes more stable with the increasing value of x (monoclinic⟶ orthorhombic). The value of the lattice parameters increase with the increase of the Lanthanum content in $YCrO_3$

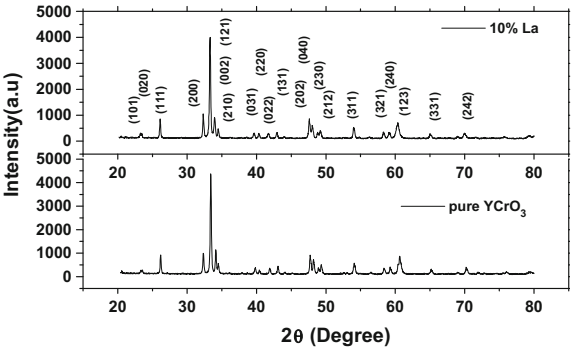


Fig. 1 XRD pattern of pure $YCrO_3$ and 10% La-doped $YCrO_3$

Table 1 The values of the lattice parameters for undoped (pure) and 10% La-doped samples

Lattice parameters	Undoped	10% La
a	7.6301	7.6556
b	7.5777	7.6175
c	7.5713	7.5853

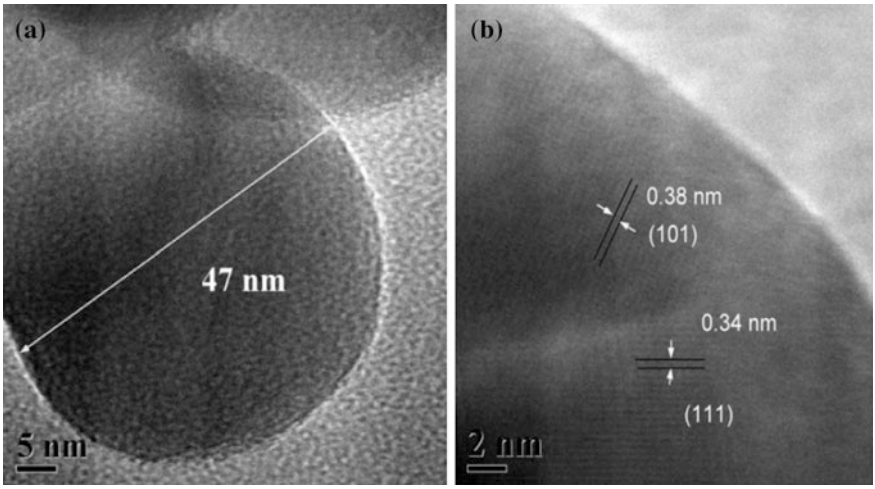


Fig. 2 **a** and **b** represents the TEM image and HR TEM image of 10% La-doped samples

compound and as a consequence lattice volume also increases as La^{3+}(1.216Å) has larger ionic radius than Y^{3+}(1.075Å). The value of the lattice parameters for undoped and 10% La-doped samples from the Rietveld analysis are listed in Table 1.

Figure 2a shows the TEM image of 10% La-doped sample. Figure 2b shows the High Resolution TEM image of 10% La-doped sample. From HRTEM image the

lattice spacings are calculated as 0.38 nm and 0.34 nm which represents (101) plane and (111) plane of $YCrO_3$ respectively, confirming the formation of Yttrium Chromite nanoparticles. From the TEM image the average particle size is estimated and it is found to be 47 nm.

3.2 Conductivity Analysis

To explain the conduction mechanism the temperature variation of conductivity (DC) is studied for doped and undoped samples within the temperature 298–523 K. It has been noticed that the conductivity (DC) increases with the increase of doping concentration for all the samples. It has been noticed that conductivity (DC) is also increasing with the increase of temperature which proves the semiconducting behavior of all the samples.

According to Arrhenius model [6] where the conductivity is explained as

$$\sigma_{dc}(T) = \sigma_0 \exp\left[-\left(\frac{E_g}{K_B T}\right)\right] \tag{2}$$

Here, σ_0 is a constant, E_a is activation energy of DC conduction, K is the Boltzmann constant. The activation energy is measured from the slope of the graph in Fig. 3 and the values are listed in Table 2 for all the samples. It is clear that the conductivity (DC) has been increased but the activation energy has been decreased for the increase of Lanthanum content of the samples, as the charge carriers require less amount of energy to be thermally activated for the hopping process for the transport mechanism.

With the increase of the La concentration the conductivity is increased and the activation energy is decreased for all the experimental samples implying the transport process is increased with the expense of the activation energy.

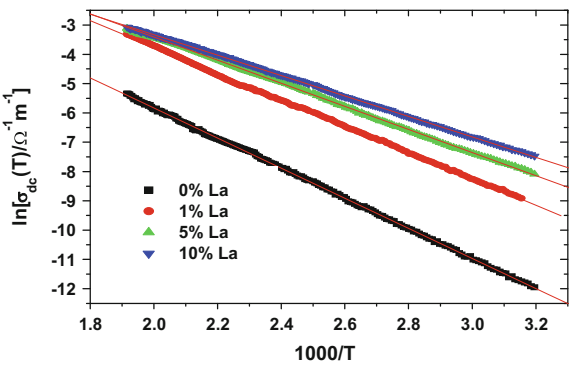


Fig. 3 Temperature dependence of conductivity (DC) for the experimental samples

Table 2 The dc activation energy for all the samples

Sample name	Dc activation energy (eV)
0% Lanthanum	0.44
1% Lanthanum	0.39
5% Lanthanum	0.34
10% Lanthanum	0.30

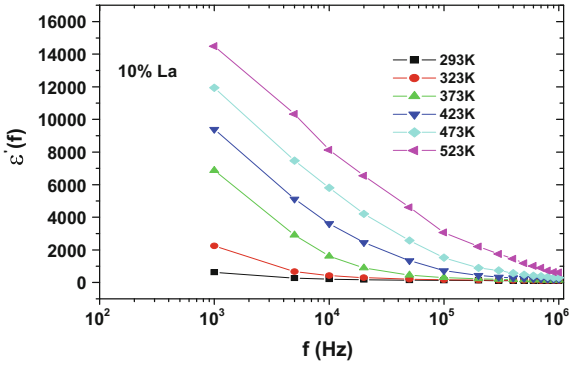


Fig. 4 The frequency variation of $\varepsilon'(f)$ of 10% La-doped sample for different temperatures

3.3 Dielectric Property Analysis

The complex dielectric constant [7] can be written as $\varepsilon(f) = \varepsilon'(f) - i\varepsilon''(f)$. The frequency dependence of real part of dielectric constant $\varepsilon'(f)$ and the imaginary part of dielectric constant $\varepsilon''(f)$ for different temperatures are expressed as follows

$$\varepsilon'(f) = \frac{C(f)d}{\varepsilon_0 A} \tag{3}$$

$$\varepsilon''(f) = \varepsilon'(f)D(f), \tag{4}$$

where C(f) represents capacitance, D(f) is the dissipation factor, ε_0 is the free space permittivity, A electrode area and d is the thickness of the experimental samples. For 10% La-doped sample the frequency variation of $\varepsilon'(f)$ and $\varepsilon''(f)$ for different temperatures are depicted in Figs. 4 and 5. Similar variation is also noticed for the remaining experimental samples. It can be concluded from the graph that both $\varepsilon'(f)$ and $\varepsilon''(f)$ have strong temperature variation when the temperature is high and the frequency is low. For lower frequency region the temperature dependence of $\varepsilon'(f)$ is greater compared to high frequency. At low frequency, the electronic, ionic, orientational and interfacial polarization [8] all are present and they gives their contribution. But for higher frequency the dipoles are not be able to rotate rapidly, so the oscillation of the dipoles cannot follow the applied field. As a result the orientational polarization is ceased. So for the interfacial and space charge polarization the dielectric constant value reduces for lager frequency [9].

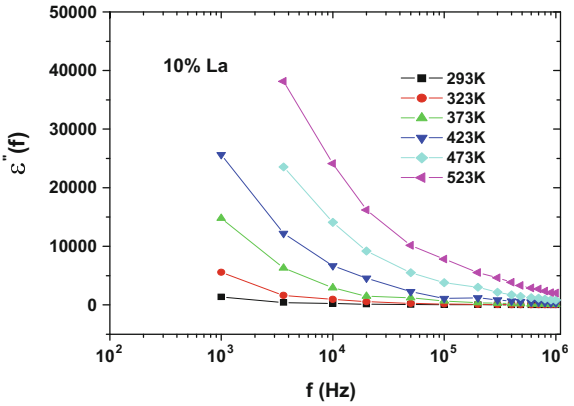


Fig. 5 The frequency variation of $\varepsilon''(f)$ for 10% La-doped sample for different temperatures

3.4 Impedance Analysis

From the Maxwell–Wegner capacitor model [10] the complex impedance may be explained with the help of an equivalent circuit with capacitance and resistance for the grain and interfacial grain boundary part. It [11–14] is written as

$$Z = \frac{1}{i\omega C_0 \varepsilon(\omega)} = Z' - iZ'' \tag{5}$$

$$Z' = \frac{R_g}{1 + (\omega R_g C_g)^2} + \frac{R_{gb}}{1 + (\omega R_{gb} C_{gb})^2} \tag{6}$$

$$Z'' = \frac{\omega R_g^2 C_g}{1 + (\omega R_g C_g)^2} + \frac{\omega R_{gb}^2 C_g}{1 + (\omega R_{gb} C_{gb})^2} \tag{7}$$

Here R_g and R_{gb} are grain and grain boundary resistances respectively. C_g and C_{gb} are the grain and grain boundary capacitances respectively. C_0 is the free space capacitance and $\omega = 2\pi f$. The real part of complex impedance has been estimated with the help of the equation

$$Z'(f) = \frac{\varepsilon''(f)}{[\omega C_0 \left(\varepsilon'(f)^2 + \varepsilon''(f)^2\right)]} \tag{8}$$

$Z'(f)$ is estimated using Eq. 8. The variation of $Z'(f)$ verses frequency (f) is shown in Fig. 6 at the temperature of 298 K. The points show the experimentally taken data and the solid line is the fitting of Eq. 8. Here R_g, R_{gb}, C_g and C_{gb} values are decreasing from pure to 10% Lanthanum doped sample and are listed in Table 3. It can be said that the grain boundary has superior contribution with respect to the grain part.

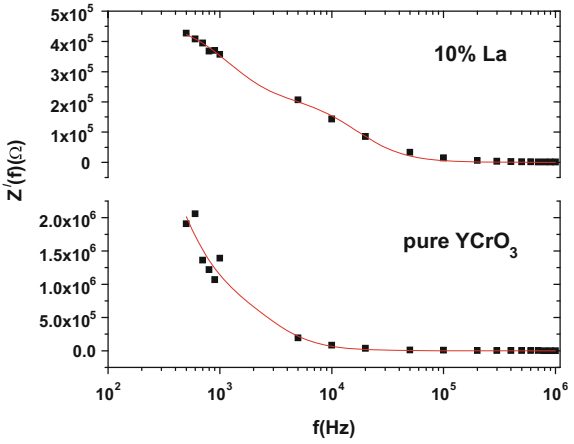


Fig. 6 The frequency variation of $Z^{'}(f)$ at room temperature for pure and 10% doped samples

Table 3 The values of R_g, R_{gb}, C_g, and C_{gb} for 0 and 10% doped samples

Sample name	R_g (MΩ)	R_{gb} (MΩ)	C_g (nF)	C_{gb} (nF)
0% Lanthanum	0.82	3.59	0.07	0.12
10% Lanthanum	0.20	0.26	0.03	0.05

4 Conclusion

The Lanthanum-doped $YCrO_3$ nanoparticles are synthesized by sol–gel method. XRD confirms the nanocrystalline nature of the samples and Rietveld analysis shows that the values of the lattice parameters and the lattice volume increase with the increase of the La concentration. The average particle size from the TEM measurement is found to be 47 nm. The DC conductivity analysis shows the semiconducting behavior of the samples and it is also increased with the increase of the La content. Activation energy reduces as the doping content increases. The frequency dependence of dielectric constant is also studied and it can be concluded that the grain boundary part governs over the grain part for resistance and capacitance.

Acknowledgements Authors are very much grateful to DAE-BRNS, Government of India and the Centre of Excellence, TEQIP-II, NIT Durgapur for the help in research work.

References

1. Weinberg, I., Larssen, P.: Nature **192**, 445 (1961)
2. Oryshich, I.V., Poryadchenko, N.E., Bega, N.D., Rakitakii, A.N.: Poroshk. Metall. **3–4**, 36 (1996)
3. Serrao, C.R., Kundu, A.K., Krupanidhi, S.B., Waghmare, U.V., Rao, C.N.R.: Phys. Rev. B **72**, 220101 (2005)
4. Chakraborty, N., Mukherjee, A., Sinha, S., Basu, S., Meikap, A.K.: Physica E **64**, 134–140 (2014)
5. Prasad, A., Basu, A., Mahata, M.K.: Chalcogenide Lett. **8**, 505–510 (2011)
6. Saha, S., Nandy, A., Meikap, A.K., Pradhan, S.K.: Mater. Res. Bull. **68**, 66–74 (2015)
7. Cao, W., Gerhardt, R.: Solid State Ionics **42**, 213 (1990)
8. Srivastava, K.K., Kumar, A., Panwar, O.S., Lakshminarayan, L.N.: J. Non-Cryst. Solids **33,** 205 (1979)
9. Frohlieh, H.: Theory of Dielectrics. Oxford University Press, London (1958)
10. McCrum, N.G., Read, B.E., Williams, G.: An Elastic and Dielectric Effects in Polymeric Solids. Wiley, New York (1967)
11. Maxwell, J.C.: A Treatise on Electricity and Magnetism, vol. 1. Oxford University Press, Oxford (1988)
12. Wagner, K.W.: Ann. Phys. (Lpz) **40**, 53 (1913)
13. Hippel, V.: Dielectrics and Waves. Wiley, New York (1954)
14. Sinha, R., Basu, S., Meikap, A.K.: Physica E **69**, 47–55 (2015)

Part III
Pattern Recognition and Machine Intelligence

Generic Object Detection Framework with Spatially Pooled Features

K. Venkatachalapathy, K. Kishore Anthuvan Sahayaraj and V. Ohmprakash

Abstract Generic Object detection technique plays an important role in the traffic surveillance and security-related issues. Research has been done over the past several years and accomplished great progresses via convolutional neural network (CNN) which has greatly enhanced the performance in image classification and object detection. This proposal is similar to the notion of R-CNN [1], presents a novel method that combines the spatially pooled features (sp-Cov) as a part of aggregated channel (ACF) and CNN for object detection. The proposed technique uses sp-Cova and ACF to select the possible object on interest regions and then trains a CNN model to filter out non-face candidates. Merging the results of sp-Cov + ACF and CNN to get the final detection window(s). The proposed framework achieves the good performance with state-of-the-art methods on numerous benchmark data sets.

Keywords Object detection · CNN · Neural networks

1 Introduction

The visual surveillance systems are the targets that are used in the real-time observations like humans or vehicles detection and also provide detailed information about the environment. Visual surveillance system has various applications and also used for security monitoring, vehicle detection, traffic flow measuring,

K. Venkatachalapathy (✉) · K. Kishore Anthuvan Sahayaraj · V. Ohmprakash
Annamalai University, Chidambaram, India
e-mail: omsumeetha@rediffmail.com

K. Kishore Anthuvan Sahayaraj
e-mail: antonio.kishore@gmail.com

V. Ohmprakash
e-mail: prakashohm96@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_11

accident detection on highways. Detection and tracking an object from a video to analyse and to predicate their behavior is processed as initial step. Due to the rapid movement of multiple objects, tracking is complex and challenging than the single object tracking. To detect the human face from an image Viola Jones et al. had proposed an object detection framework. This technique is adaptive boosting that is used for speeding a binary classifier and also applied in machine learning to create real-time detectors.

Object detection progress has been gradually increased in recent years. The detection problem R-CNN treats by object proposal windows, in CNN architecture it uses R-CNN to improve sharing computation across overlapping windows by adding Region Of Interest (ROI). This method is used for counting object by hyperparameter tuning and also carried out by other detection method by these approaches can also count the instances of an object instead of the exact localization by this also have an ability to count the object in modifying existing models to trade-off the ability to localize which is a novel finding for deep models.

In this proposal applying the object categorization method for large object it can divide into several subcategories and train one sub-detector for each subcategory or else train in a single detector for whole object class, before integrating them together all should be tested the raw detection results subdetectors to eliminate the non-maximum suppression are used to redundant bounding boxes. Hence a framework employs detector of different classes, detection result is done carefully.

2 Related Work

Generic Object Detection: Object detection is an interesting but important application in the computer vision community. It has reached successful outcomes in many practical applications such as face detection and pedestrian detection [2–5]. Complete analysis of object detection can be found in [3–6]. This section briefly analyses several generic object detection methods.

One standard object detector is the detection framework of Viola and Jones which uses a sliding window search to achieve effective classification and precise location [4]. The other commonly used framework is using a linear support vector machine (SVM) classifier with histogram of oriented gradient (HOG) features, which has been employed successfully in pedestrian detection [3]. These frameworks attain splendid detection results on rigid object classes. However, for object classes with large intraclass dissimilarity, their detection performance falls down dramatically [7]. In order to deal with appearance variations in object detection, a deformable parts model (DPM) based method has been suggested by Felzenszwalb et al. [6]. This method depends on a variant of HOG features and window template matching, but explicitly models deformations using a latent SVM classifier. It has been employed successfully in many object detection applications [8–10]. In addition to the DPM, visual subcategorization [11] is another common approach to improve the overview performance of detection model. It divides the whole object

class into multiple subclasses such that objects with similar graphical appearance are clustered together. A sub-detector is trained for each subclass and detection results from all sub-detectors are merged to generate the final results. Recently, a new detection framework which uses aggregated channel features (ACF) and an AdaBoost classifier has been proposed in [12]. This framework uses complete sliding window search to detect objects at multi-scales. It has been employed successfully for many practical applications [7, 13, 14].

3 Our Approach

The conventional sliding window-based method of Viola Jones is still the most successful and practical object detector. The VJ framework has two main components, a dense feature extractor and a cascade classifier. In this proposal we propose a common object detection framework for traffic scene view based on the VJ framework, but framework can work on a number of different classifiers to detect target of different classes. Apart from this an object subclassification method is used to improve the performance and employ spatially pooled features to enhance robustness and efficiency.

The Fig. 1 shows the overview of a novel framework. In the training phase we check for intraclass variations of input object class with respect to object properties example size, orientation, aspect ratio and occlusion. If the variations are large the object subclassification method is applied to categorize the objects of different class into multiple subcategorize and train one sub-detector for each segmentation. Otherwise we train a single detector for the entire object class. In the testing phase the detected results from all subcategories are merged together for score calibration. The Associative Subitizing is applied to count the objects in the image by dividing the image into regions. The non-maximum suppression is used to eliminate redundant bounding boxes.

3.1 Object Subclassification

The variations of the object image can change significantly as viewpoint changes. The conventional method of VJ framework cannot tackle these variations. The subclassification method applies an unsupervised clustering method to one specific features space of the training samples to generate multiple subcategories.

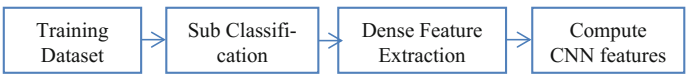


Fig. 1 Object detection system overview

This simplifies the original learning problem by divide and conquer method. The HOG and ACF are used for the clustering method. The HOG method neglects the color information and captures only the shapes of the objects. An ACF method combines both the color and gradient information. As mentioned in [15] total of 10 features channels are used for clustering. All samples are resized to the median object size in the training. The clustering method is used to generate a predefined number of clusters on a feature space. Traditional clustering schemes, suffer from the cluster degeneration which means that a few clusters claim most data samples [15]. The cluster degeneration problem can be improved by using spectral clustering. Spectral clustering followed by k-means often overtakes the traditional schemes. Implementation the normalized spectral clustering using the algorithm proposed in [16].

3.2 Feature Extraction

The framework introduces spatially pooled features as part of aggregate channel features in the training phase to generate a dense feature. All dense feature channels are aggregated in 4 × 4 blocks in order to return fast pixel lookup features.

Aggregated Channel Features (ACF): Aggregated channel features are extracted from multiple image channels using pixel lookup method. Many image channels are available for extracting dense features. The dense features are extraced only once in the evaluation phase so the image need not be transformed seperately.

Spatially Pooled LBP: Spatial pooling is used to merge multiple visual descriptors acquired at nearby locations into a lower dimensional descriptor over the pooling region. Tracking the work of [7] which is shown that pooling can enhance the strength of two hand-crafted low-level features, covariance features [17] and LBP [2]. A covariance matrix is a positive semidefinite matrix which provides a degree of the relationship between multiple sets of variants. The diagonal elements of a covariance matrix represent the modification of each feature and non-diagonal elements represent the relationship between different features.

The robustness and spatial invariance of the covariance descriptors can be improved by applying sharing method. Max sharing is used in this framework as it has been shown to overtake average sharing in image classification [18, 19]. Max sharing uses the maximum value of a sharing region to represent the shared features in the region. It aims to hold the most relevant information and discard irrelevant details and noises over the sharing region. Similar to the sp-Cov features, divided the image into multiple dense areas and LBP histogram is calculated over pixels within each area. In order to improve the invariance to image distortion and translation, performing max sharing to represent the LBP histogram in the pooling region. The shared features extracted from each sharing region are called the spatially pooled LBP (sp-LBP) features in [7].

Implementation: To extract LBP in the proposal process, extraction from a 4 × 4 pixel patch on the 33—neighborhood. We extract the 58-dimension LBP

histogram. In this work, the patch step size, the sharing region, and the sharing spacing step are set to 1 pixel, 4 × 4 pixels, and 2 pixels, respectively. As an alternative of extracting LBP histograms from multi-scale patches, the sp-LBP and LBP are pooled as channel features.

3.3 Compute CNN Features with Dense Features

Once dense features have been extracted to train a classifier process is carried out. The unsupervised pre training process is done with the large auxiliary dataset using image annotations only. Then the supervised fine tuning is done to increase the mAP performance. The CNN can be used a black box feature extractor, this gives better results in several applications like scene classification, fine grained sub categorization and domain adaptation. To adapt CNN to the new task and the new object class, we continuation of processed HOG training of the CNN parameters using only wrapped regions proposals.

3.4 Online Training

To improve the detection performance some methods are applied as a post processing step.

Calibration of scores: The calibrated score is defined as

$$g_i^k = \frac{1}{1 + \exp(A_k.s_i^k + B_k)} \tag{1}$$

If the features are extracted from multiple sub detectors, the results have to be merged to get the result. For sample i in subcategory k, its confidence score is the output of the ensemble classifier which is defined as

$$s_i^k = \sum_{t=1}^{T} a_t h_t(x_i^{-k}) \tag{2}$$

Non-maximum Suppression (NMS): The main aim of the NMS is to reduce the redundant bounding boxes among the detection results. When multiple bounding boxes overlap each other, NMS will remove the lower scored detection and retain the high scored detection. The overlap ratio is defined as

$$a_0 = \frac{area(B_1 \cap B_2)}{area(B_1 \cup B_2)} \tag{3}$$

where B_1 and B_2 are two different bounding boxes.

Table 1 Evaluation of the overall run time of our framework

Feature Combinations	Feature extraction [s]	Cars detection [s]	Total runtime [s]
ACF	0.10	0.40	0.65
Sp-LBP +ACF	0.35	1.20	1.95
Sp-Cov + ACF	5.50	1.30	7.20
Sp-Cov + sp-LBP + ACF	5.75	1.75	8.00

Fusion of Detection Results: Our framework can detect generic objects simultaneously using detectors or sub-detectors of different classes. Since our framework uses multiple sub-detectors to detect objects NMS usually eliminates the redundant bounding box. Consider a scenario where a car is occluded by a cyclist. If their overlap ratio exceeds the threshold, NMS will remove the lower scores detection. To overcome this problem a fusion process is carried out to merge all the detected results. The NMS is applied at each individual class separately to filter out redundant bounding boxes from different classes. Thus by keeping the true positive detected objects in different classes.

4 Experiments

In this section an experiment on KITTI dataset. The estimation benchmark consists of nearly 7400 training images and 7500 test images, comprising a total of 80,000 labeled objects of annotated objects in traffic scene. All experiments are carried out by combining all the features (ACF + sp-Con + sp-LBP).The average runtime of each component of the framework can be seen in Table 1. For feature extraction we observe that ACF features can be extracted quickly when compared with other features. The sp-Con features extracting increases as the number of sub detector increases, thus overall time also increased. Finally, we observe that the detection results using ACF + sp-LBp features can provide better results between detection performance and system runtime.

5 Conclusion

This proposal presented sharing spatially pooled convolutional features with the detection network; the region proposal step is nearly cost-free. Proposed method enables a unified, deep-learning-based object detection system. The learned RPN also expands region proposal quality and the general object detection accuracy.

References

1. Girshick, R.B., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of IEEE Conference on Computer Vision Pattern Recognition, pp. 580–587 (2014)
2. Ahonen, T., Hadid, A., Pietikäinen, M.: Face recognition with local binary patterns. In: Proceedings of European Conference on Computer Vision, pp. 469–481 (2004)
3. N. Dalal and B. Triggs.: Histograms of oriented gradients for human detection. In: Proceedings of IEEE Conference on Computer Vision Pattern Recognition, pp. 886–893 (2005)
4. Viola, P., Jones, M.J.: Robust real-time face detection. Int. J. Comput. Vis. **57**(2), 137–154 (2004)
5. Wang, X., Han, T.X., Yan, S.: An HOG-LBP human detector with partial occlusion handling. In: Proceedings of IEEE International Conference on Computer Vision, pp. 32–39 (2009)
6. Felzenszwalb, P.F., Girshick, R.B., McAllester, D.A., Ramanan, D.: Object detection with discriminatively trained part-based models. IEEE Trans. Pattern Anal. Mach. Intell. **32**(9), 1627–1645 (2009)
7. Paisitkriangkrai, S., Shen, C., van den Hengel, A.: Strengthening the effectiveness of pedestrian detection with spatially pooled features. In: Proceedings of European Conference on Computer Vision, pp. 546–561 (2014)
8. Geiger, A., Wojek, C., Urtasun, R.: Joint 3d estimation of objects and scene layout. In: Proceedings of Advances in Neural Information Processing Systems, pp. 1467–1475 (2011)
9. Torres, J.J.Y., Bergasa, L. M., Arroyo, R., Lazaro, A.: Supervised learning and evaluation of kitti's cars detector with DPM. In: Proceedings of IEEE Intelligent Vehicles Symposium, pp. 768–773 (2014)
10. Yan, J., Zhang, X., Lei, Z., Liao, S., Li, S.Z.: Robust multi-resolution pedestrian detection in traffic scenes. In: Proceedings of IEEE Conference on Computer Vision Pattern Recognition, pp. 3033–3040 (2013)
11. Divvala, S.K., Efros, A.A., Hebert, M.: How important are deformable parts in the deformable parts model? In: Proceedings of European Conference on Computer Vision Workshop, pp. 31–4 (2012)
12. Dollár, P., Appel, R., Belongie, S., Perona, P.: Fast feature pyramids for object detection. IEEE Trans. Pattern Anal. Mach. Intell. **36**(8), 1532–1545 (2014)
13. Mathias, M., Timofte, R., Benenson, R., Gool, L.J.V.: Traffic sign recognition-how far are we from the solution? In: Proceedings of International Joint Conference on Neural Networks, pp. 1–8 (2013)
14. Ohn-Bar, E., Trivedi, M.M.: Fast and robust object detection usingvisual subcategories. In: Proceedings of IEEE Conference on Computer Vision Pattern Recognition Workshop, pp. 179–184 (2014)
15. Jain, A.K., Murty, M.N., Flynn, P.J.: Data clustering: a review. ACM Comput. Surv. **31**(3), 264–323 (1999)
16. Ng, A.Y., Jordan, M.I., Weiss, Y.: On spectral clustering: analysis and an algorithm. Proc. Adv. Neural Inf. Process. Syst. **2**, 849–856 (2002)
17. Tuzel, O., Porikli, F., Meer, P.: Region covariance: a fast descriptor for detection and classification. In: Proceedings of European Conference on Computer Vision, pp. 589–600 (2006)
18. Coates, A., Ng, A.Y.: The importance of encoding versus training with sparse coding and vector quantization. In: Proceedings of Interantional Conference on Machine Learning, pp. 921–928 (2011)
19. Li, B., Wu, T., Zhu, S.: Integrating context and occlusion for car detection by hierarchical and-or model. In: Proceedings of European Conference on Computer Vision, pp. 652–667 (2014)

Hybrid PSACGA Algorithm for Job Scheduling to Minimize Makespan in Heterogeneous Grids

Amit Chhabra and Oshin

Abstract Grid provides a clear, coordinated, consistent and reliable computing medium to solve complex sequential and parallel applications through the use of idle CPU cycles. Scheduling optimizes the objective function(s) by mapping the parallel jobs to the available resources. Owing to the heterogeneity of resources in grid, scheduling associates with class of NP-hard problems due to which reaching optimal solution surpasses the time constraint. Metaheuristic algorithms take polynomial time to reach the near-optimal solutions for NP-hard problems. Major research issues in metaheuristic algorithms are solution quality and convergence speed that have been revised by using consolidation approach. This paper proposes a hybrid PSACGA algorithm that consolidate the features of Particle Swarm Optimization (PSO), Ant Colony optimization (ACO) and operators of Genetic algorithm to solve parallel job scheduling problem. Experimental results of the proposed technique are compared with existing deterministic and metaheuristic job scheduling algorithms. Experimental results have indicated that the proposed hybrid PSAGA algorithm provides better performance than existing contemporary algorithms.

Keywords Grid · Job scheduling · Ant colony optimization · Particle swarm optimization and genetic algorithm

A. Chhabra (✉) · Oshin
Department of Computer Engineering and Technology, Guru Nanak Dev University, Amritsar, India
e-mail: amit.cse@gndu.ac.in

Oshin
e-mail: eroshinbhardwaj@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_12

1 Introduction

Grid computing has offered a vast platform for virtual system to distribute their services over the network of resources [1]. It aimed at reducing cost in addition to turnaround time as well as achieving substantial program throughput [2]. Grid computing [3, 4] provides a platform over Internet to researchers and organizations to solve complex problems by integrating and sharing computational power of geographically distributed commodity clusters. Based on offered utilities including calculation, facts, intercommunication, service, knowledge and application support, grid types [5, 6] can be classified into the utility grid, knowledge grid, data-centric grid, computational grid. Grid computing has been effectively put in place to remedy numerous real-life complications [7, 8]. For any grid-based computing organization, scheduling problems are a major concern [9]. A computational system is majorly influenced by scheduling algorithm [10]. Elementary approaches like greedy algorithm or FCFS can be used for the implementation of the scheduling algorithm. However, a grid system consisting of heterogeneous computational resources requires more matured algorithm to obtain better results. Job scheduling in a grid environment is a known NP-complete problem [11]. Heuristic and metaheuristic algorithms are often used to compromise with the complications of these problems [12] and develop optimum or sometimes near-optimum solutions. Primary heuristic technique suggested for scheduling are Tabu Search, Simulated Annealing and Swarm Intelligence with two most opted approaches ACO and PSO.

ACO is simulated from shortest route seeking nature of ant species. Limiting makespan/flow-time along with the economic price is the intended direction. ACO technique has slow convergence speed and poor scattering ability [13]. The problem becomes more severe with the increase in problem size. PSO has many components familiar to Genetic Algorithm. PSO is inspired from the co-operative and competitive nature of swarm uses star and ring topology that scales from global to more local optimization. PSO has balanced ratio of exploration and exploitation mechanism. Parallel exploration of neighborhood in solution space makes PSO less prone to stuck in local minima. GA shows uncertainty in local search and repeatedly gets stuck in local minima. Unlike GA, PSO does not have crossover and mutation operator and the information is shared only by ‘gbest’ particle whereas in GA, entire population of particles is inclined to reach the best solution. The proposed hybrid PSACGA algorithm tends to overcome the limitations of individual ACO, PSO and GA metaheuristic algorithms and experimental results prove its worth by optimizing makespan and flow time of the parallel job scheduling problem.

2 Proposed Method

2.1 Particle Swarm Optimization

Kennedy and Eberhart proposed PARTICLE SWARM OPTIMIZATION—Stochastic, population-based technique animated from the behaviour of flocks of birds [14]. Each individual is analogous to particle that escalates over multi-dimensional search space. Each particle is randomly initialized with certain velocity and position. In each iteration, particle updates its position based on personal experience and its neighbor's experience, whereas velocity is updated according to its best position achieved so far and best position achieved by any particle in entire population [15]. PSO attempts to equate exploration and exploitation rate by combining local with global search [16]. Initial step involves encoding of problem. Kousalya [17] used n * m matrix where n is no. of servers and m is no. of tasks. 0 and 1 can be only accommodated in matrix where one column can have only one '1' value. Job allocation is specified by columns and resource allocation is specified by rows. Matrix is used to represent velocity. Different approaches are being used to initialise population. Assignment of Shortest job first to processor with fastest speed [18] can be used to get better solution [19]. Zhang [20] applied VNS that promotes local search and increases exploitation capability of PSO. To get balance between the speed of machine and length of jobs PSO-integer method is effective [21]. Crossover and mutation operation of Genetic algorithm are used to improve exploration and exploitation rate and to get improved speed [22].

Use of PSO in Job Scheduling

(a) **Assigning Position to particle**

Initial step for implementing PSO is to represent mapping between particles and problem space. We encode position in the form of n * m matrix where m number of jobs is mapped to n number of servers. Let X_i be the position matrix of ith particle. X_i (j, k) signifies that kth job is executed by jth server.

(b) **Velocity of particle**

n * m matrix represents velocity possessed by each particle and each value lies in the range [−Vmax, Vmax]. Let Vi be the velocity of ith particle then

$$\mathrm{Vk}(\mathrm{j},\mathrm{k}) \in [-\mathrm{Vmax}, \mathrm{Vmax}](\forall \mathrm{j},\mathrm{k}), \mathrm{j} \in \{1, 2 \ldots \mathrm{m}\},$$

(c) $\mathrm{k} \in \{1, 2 \ldots \mathrm{n}\}$

(d) ***Pbest* and *gbest***

'Pbest' and 'gbest' are n * m matrices taken as memory component. Pbest is best position particle has attained personally and gbest is best position attained by any particle in entire population. Let the best position accomplished by ith particle is pbesti and best position accomplished by entire population is gbesti. Update Pbest and gbest in each step. Firstly evaluate fitness value corresponding to each particle (let X_i)

$$\begin{aligned} &If\ Xi > pbesti \\ &\qquad \{pbesti = Xi\} \\ &If\ pbesti > gbesti \\ &\qquad \{gbesti = pbesti\} \end{aligned}$$

(e) **Updating particle**

Velocity and position of each particle updated by Eqs. 1 and 2 respectively:

$$\begin{aligned} &V_i^{(s+1)}(j,k) = \\ &w.V_i^s(j,k) + a_1 rand_1\left(pbest_i^s(j,k) - X_i^s(j,k)\right) + a_2 rand_2\left(gbest_i^s(j,k) - \right. \\ &\left. X_i^s(j,k)\right) \end{aligned} \tag{1}$$

$$X_i^{(S+1)}(j,k) = X_i^s(j,k) + V_i^{s+1}(j,k) \tag{2}$$

$V_i^s(j,k)$ is the value in the *jth* row and *kth* column of *ith* velocity vector in *sth* time stamp, a_1 and a_2 are acceleration constant, $rand_1$ and $rand_2$ lies in interval [0, 1] are random values, w is inertia weight control the exploration and exploitation rate, $X_i^s(j,k)$ is *jth* row and *kth* column of *ith* position vector in *sth* time stamp.

2.2 Ant Colony Optimization

Marco Dorrigo designed "nature inspired" algorithm termed as ACO, encouraged from the forging role of ants which produces an optimum or near-optimum path from their dwell to food [22, 23]. It had been discovered that ants when searching for meals, are generally very motivated by pheromones. Discrete optimization problems implement ACO to reach satisfactory results. ACO solves various problems such as Knapsack problem, TSP [24], quadratic assignment problem, job scheduling in cloud and grid environment, Job shop scheduling [25]. Ku Ruhana, Ku-Mahamud combines Ant Colony System (ACS) and Max–Min to schedule jobs in grid computing. Resource availability is checked by updating the status in matrix. Resource table of grid is checked by using the concept of agent in the algorithm [26, 27]. Tawfeek [28] focused that each ant will visit each of the machine once only and fitness is evaluated on the basis on execution time. The results showed that

with the increase in the number of jobs to be executed, ACO performs better in terms of time that FCFS and RR (Round Robin). Makespan is considered the objective function. Bagherzadeh and MadadyarAdeh [29] used deterministic approach initially and used its results to initialize biased ants. Heuristic information, pheromone and execution time of job is considered. Bagherzadeh and Madadyar-Adeh [29] used deterministic approach to initialize biased ants. Pheromone, Heuristic information and job standard deviation and execution time are used to execute jobs. Liu [30] changes value of pheromone value for scheduling tasks. Evaporation rate keeps on changing but is never zero. Ants [31] are allocated resources using probability function and they start executing tasks. Mapping of task to resource is based on precedence constraint.

ACO in Job Scheduling

(a) **Initialization**

Initial step for implementing ACO is mapping of ants to problem solution and represented as matrix. Scheduling of independent jobs requires less number of ants or ants should be equal to number of jobs. Pheromone value is initialized with a positive constant.

(b) **Solution Construction**

Probability function is used to determine next job and resource on which it will execute [31].

$$pk_{j,k} = \frac{\tau_{j,k}^{\alpha} * \eta_{j,k}^{\beta}}{\sum \tau_{j,k}^{\alpha} * \eta_{j,k}^{\beta}} \tag{3}$$

Where α signifies impact of pheromone value, β signifies impact of heuristic information, $\eta_{j,k}$ is the heuristic information, $\tau_{j,k}$ is the pheromone trail linked to job j_j and resource r_k. It is repeated until all the jobs have been allocated to resources and ants have built complete solution matrix.

(c) **Pheromone update**

While constructing the vector, every ant traverses and locally updates the quantity of pheromone using the following rule [32]:

$$\tau_{j,k} = (1 - r).\tau_{j,k} + r.\Delta\tau_{j,k} \tag{4}$$

Where r is the parameter that evaporates pheromone and $0 < r < 1$. After all the jobs have been allocated, the pheromone is globally updated. The shortest path covering ant is only permitted to append the pheromone. The rule for globally updating the pheromone is as follows:

$$\tau_{j,k} = (1-r).\tau_{j,k} + \beta.\Delta\tau_{j,k} \tag{5}$$

where,

$$\Delta\tau_{j,k} = \begin{cases} (l_{bb})^{-1} \\ 0, \ otherwise \end{cases}, \ if \ (e,f) \in bb^{best}$$

l_{bb} is the extent of the global best (bb^{best}) route. The ant solution that minimize makespan/flow time is the best solution achieved by the ant.

2.3 Genetic Algorithm

Genetic algorithm is an evolutionary algorithm based on "SERVIVAL OF FITEST" concept [33]. Chromosomes are set of genes and represented as set of solutions to the given problem. The algorithm starts by randomly generating population [33]. Fitness function evaluates the competency of chromosome for particular problem domain. Crossover and mutation operators are applied on chromosomes preferred through fitness function to produce new and better population.

Use of GA in Job Scheduling

In scheduling, chromosomes (ch) can be represented by 1-Dimensional vector or 2-Dimensional vector. In 1-Dimensional vector, ch is vector of size m, where m is the number of jobs and resource allocated to jth job is represented by ch(j). In 2-Dimensional vector, order of jobs is represented by one of the dimension and other dimension represents resource to which job is allocated. Fitness function can include flow time, makespan and execution time. Different crossover operators have been by suggested different authors such as One-point, Two-point, ordered crossover, Uniform crossover, Cycle crossover, Fitness based and partially matched crossover. Likewise different mutation operators are also used such as Swap, Move, Move, Swap, etc.

2.4 Proposed Hybrid PSACGA Algorithm

The proposed technique consolidates the features of PSO, ACO and mutation and crossover operator of GA. The hybridization overcomes the gaps of stand-alone algorithms. Fitness is evaluated on the basis of makespan. Instead of being highly skilful in solving optimization problem, it has slow convergence speed and poor scattering ability. The problem is more severe with the increase in problem size. PSO has many components familiar to Genetic Algorithm—the Evolutionary

```
   1. Initialize
Number of jobs, number of servers, execution time (EX)
   2. While number of jobs/solutions do
      Populate solution vector  with random id and servers
                  While length of elements of solution vector  do
                              While length of elements of solution vector do
                                  If solution vector id equals jobs id, Then
                                 Assign burst time of job to solution vector
                                  End if
                               End while
                  End while
       End while
   3. While no. of  servers do
      Populate EX
      End while
//PSO PHASE BEGINS
   4. Map jobs to the respective servers to obtain k number of random
      schedules.
      Initialize

   x = solution id, y = server id, n = job iteration, NC =no. of iterations,
   r=evaporation factor, a, b = constant, m = memory length, velocity vector
   and vector
   5. Evaluate the schedules based on optimization metric Makespan, Calculate
      the distance using equation (3.4.1) for each schedule
   6. For 1:k
      For 1:m
   While (f(k)<F(m) )
                a) Update pcbest vector and gbest vector.
                b) Update velocity and position using equation (3.4.2) and
                   (3.4.3) respectively.
End while
      End For
      End for
// PSO ENDS
// ACO PHASE BEGINS
   7. Using PSO information evaluate pheromone matrix.
   8. For 1: m do
             Calculate probability of next job using: Equation (3.4.5)
              End for
   9. While job do
                      If random number<probability than
                              Move to next job
                       Else
                            Update probability
                      End if
End While
   10.      Obtain schedule.
// BEGIN GA OPERATORS
   11.      For 1: m do
Perform crossover and mutation to obtain schedule.
End for
// END GA OPERATORS
   12.      If f(schedule1)>f(schedule)
                    Best Schedule= Schedule1
               Else
                    Best Schedule= Schedule
            End If
   13.      For1: m do
                     Update pheromone using: Equation (3.4.6).
End for
// END ACO
   14.      Update solution Vector and update EX.
   15.      Calculate mean waiting time, mean flow time and Makespan.
```

Fig. 1 Pseudo code of proposed algorithm

Technique. PSO is inspired from the co-operative and competitive nature of swarm uses star and ring topology that scales from global to more local optimization [34, 35]. PSO has balanced ratio of exploration and exploitation mechanism. Parallel exploration of neighbourhood in solution space makes PSO less prone to stuck in local minima. GA shows uncertainty in local search and repeatedly gets stuck in local minima. Unlike GA, PSO does not have crossover and mutation operator and information is shared only by *gbest* particle, whereas in GA entire population of particles is inclined to reach the best solution (Fig. 1).

Initially, PSO maps the randomly generated jobs to the servers and represented as n * m matrix where mth job is assigned to nth server. Distance between the jobs is calculated using the equation:

$$d_{ij} = \sqrt{\{(x_i - x_j)^2 + (y_i - y_j)^2\}} \tag{6}$$

The equation will calculate distance between x_i job executed on y_i and x_j job executed on y_j serve. *Pcbest* and *gcbest* is obtained on the basis on basis on minimum makespan. The pheromone matrix is initialized by PSO.

The probability function selects the next job to be allocated to which the server. Not always the pheromone value is used to search for optimal schedule; it can be done on the basis of random value selected. The output schedule of ACO and *est*, *gcbest* form PSO are combined using crossover and mutation operators. Crossover is used to produce offspring of new population that is supposed to have better fitness value. In this paper, ordered crossover is used (Fig. 2).

After crossover, mutation is the last step that randomly selects one or two position and flips their values. This study evaluates makespan, flow time and

Fig. 2 Pseudo code of crossover operator

Parents $y_1 = [y_{1,1}, y_{1,2} \ldots . y_{1,m}]$ *and* $y_2 = [y_{2,1}, y_{2,2} \ldots . y_{2,m}]$

.....

Initialization

a) Initialize children
$x_1 = [x_{1,1}, x_{1,2} \ldots \ldots x_{1,m}]$ and $x_2 = [x_{2,1}, x_{2,2} \ldots \ldots x_{2,m}]$

b) Let a and b are two crossover points
$1 \le a \le b \le m;\ h_1 = h_2 = l = b + 1;$

$\boldsymbol{j = 1;}$

Repeat

If $y_{1,j} \notin [y_{2,a}, \ldots y_{2,b}]$ *then* $x_{1,h_1} = x_{1,l}\ ;\ h_1 + +;$

If $y_{2,j} \notin [y_{1,a}, \ldots y_{1,b}]$ *then* $x_{2,h_1} = x_{2,l}\ ;\ h_2 + +;$

Until $\boldsymbol{j \le m}$

$x_1 = [x_{1,1} \ldots . x_{1,a-1}\, y_{2,a} \ldots . . y_{1,b}\, x_{1,a \ldots \ldots . . x_{1,m-a}}];$

$x_2 = [x_{2,1} \ldots . x_{2,a-1}\, y_{2,a} \ldots . . y_{2,b}\, x_{2,a \ldots \ldots . . x_{2,m-a}}];$

average waiting time and objective function is corresponding to makespan. The best schedule achieved will be having minimum makespan, that is there will be optimal mapping of jobs to the servers. The main objective of the proposed algorithm is to minimize the value of makespan.

$$\text{Min Makespan: } \min Sch_{i\ \varepsilon\ schedule}\left\{max_{j\ \varepsilon\ jobs} T_j\right\}$$

'Jobs' is set of jobs that are to be scheduled, Schedule is possible set of schedules and Tj is the time taken by jth job to finish its execution. Flow time is another criteria opted by algorithm. Flow time is sum of finishing time of all jobs.

$$\text{Flow Time} = \min\ Sch_{i\ \varepsilon\ schedule}\left\{\Sigma_{j\ \varepsilon\ jobs} T_j\right\}$$

3 Simulation Results

The experiments have been executed in a simulator which is developed in MATLAB framework. In the proposed technique, n no. of independent rigid parallel jobs is to be processed by m no. of non-identical parallel servers. Let b_{ij} is the burst time of *jth* job to be processed on *ith* server where j = 1, 2, 3, …n and i = 1, 2, 3, …m, $\{r_1, r_2, r_3 \ldots r_n\}$ be the resource requirement of set of n jobs and $\{s_1, s_2, s_3 \ldots s_m\}$ be the no. of resources possessed by m no. of servers. It is assumed that servers are divides into three classes $\{C_1, C_2, C_3\}$ with computation speed {1, 2, 3}. All the jobs are arriving at zero time and execute till the end without any intervention with burst time b_{ij}. One server can execute only one job at single time. The **PSACGA** algorithm is compared with Ant Colony Algorithm, First Come First Serve with Fastest Fit (FCFS-FF), Shortest Job First with Fastest Fit (SJF-FF), and Longest Job First with Fastest Fit (LJF-FF).

First Come First Serve with Fastest Fit (FCFS-FF): FCFS is the simplest technique opted for scheduling. The jobs are allocated resources depending on the arrival time. The one with least arrival time will get the resources first. FCFS when combined with Fastest Fit that ensures that the server with highest speed is allocated first. Jobs are processed according to their arrival time and prove unfair to the jobs that arrive late but possess least burst time. Therefore suffers from CONVOY EFFECT. Average waiting time for FCFS is quite high.

Shortest Job First with Fastest Fit (SJF-FF): In SJF, Job with least Burst Time is first allocated CPU or when the CPU is free it is first assigned to the job with minimum burst time. SJF overcomes the problem of FCFS since jobs with shortest burst time has not to wait too long for CPU. SJF combined with fastest fit execute the arriving job with minimum burst time on the server with fastest speed.

Longest Job First with Fastest Fit (LJF-FF): LJF allocates CPU to the Job with maximum Burst Time or when the CPU frees it is first assigned to the job

with highest burst time [3]. Fastest Fit ensures server with highest computation speed is assigned first.

The parameter values used by PSO-ACO-GA algorithm are in Table 1. The algorithm is run 10 times to calculate average meantime, average flow time and average waiting time.

From Tables 2, 3, 4, it is clear that the proposed PSACGA algorithm has minimum makespan, mean flow time and mean waiting time than other deterministic and metaheuristic algorithms. Figs. 3, 4 and 5 shows the arithmetic mean of makespan, average flow time and average waiting time respectively under various combinations of jobs and servers as shown in Table 4 and it can be clearly seen that PSACGA out performs other algorithms.

Table 1 Parameter value

Evaporation rate (r)	α	β	M	No. of iteration (NC)	No. of Ants (N)
0.9	1.5	2	30	15	10

Table 2 Makespan

Number of (jobs, servers)	PSACGA	ACO	FCFS-FF	SJFS-FF	LJFS-FF
(100, 10)	792	799	1065	881	920
(200, 20)	767	829	1005	913	1004
(300, 30)	803	852	1080	1050	1056
(400, 40)	815	823	1164	1441	1227
(500, 50)	804	825	1230	1101	1135
(600, 60)	835	836	1130	1093	1496
(700, 70)	831	879	1609	1277	1230
(800, 80)	803	805	1515	1151	1157
(900, 90)	844	854	1344	1328	1408
(1000, 100)	848	849	1321	1142	1405

Table 3 Average flow time

Number of (jobs, servers)	PSACGA	ACO	FCFS-FF	SJFS-FF	LJFS-FF
(100, 10)	285	283	289	235	281
(200, 20)	259	261	267	262	303
(300, 30)	259	266	269	268	270
(400, 40)	229	250	277	265	320
(500, 50)	237	236	284	240	289
(600, 60)	228	226	245	230	286
(700, 70)	256	254	286	258	309
(800, 80)	238	239	267	250	290
(900, 90)	269	270	316	280	325
(1000, 100)	253	252	283	260	301

Table 4 Average waiting time

Number of (jobs, servers)	PSACGA	ACO	FCFS-FF	SJFS-FF	LJFS-FF
(100, 10)	792	799	1065	881	920
(200, 20)	767	829	1005	913	1004
(300, 30)	803	852	1080	1050	1056
(400, 40)	815	823	1164	1441	1227
(500, 50)	804	825	1230	1101	1135
(600, 60)	835	836	1130	1093	1496
(700, 70)	831	879	1609	1277	1230
(800, 80)	803	805	1515	1151	1157
(900, 90)	844	854	1344	1328	1408
(1000, 100)	848	849	1321	1142	1405

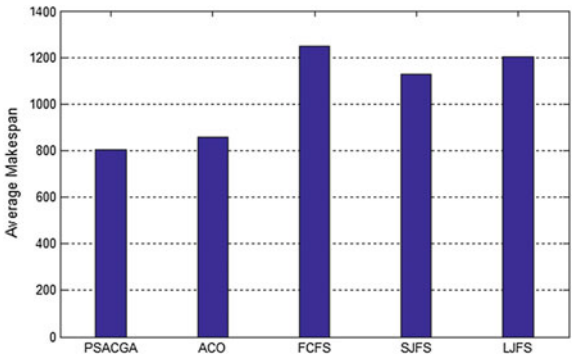


Fig. 3 Arithmetic mean of makespan

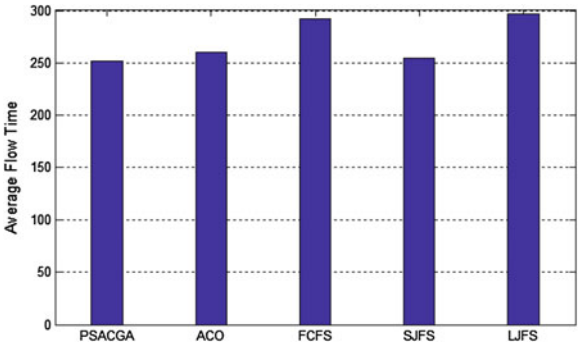


Fig. 4 Arithmetic mean of average flow time

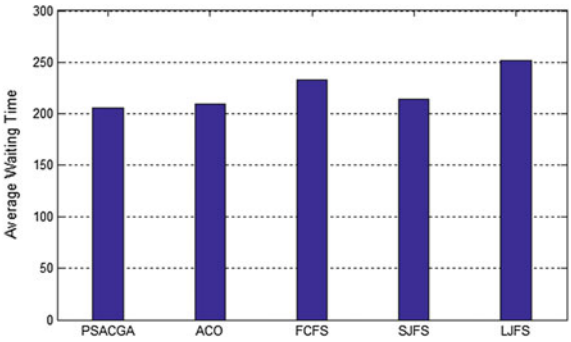


Fig. 5 Arithmetic mean of average waiting time

4 Conclusion

This paper presented a hybrid PSACGA algorithm for scheduling parallel jobs in a grid consisting of heterogeneous resources. The proposed PSACGA algorithm minimized makespan and mean flow time. PSACGA overcomes the gaps of ACO, PSO and GA techniques and able to provide better results. It has also managed to improve the convergence speed which was a major issue in ACO and GA.

References

1. Qureshi, M.B., Dehnavi, M.M., Min-Allah, N., Qureshi, M.S., Hussain, H., Rentifis, I., Tziritas, N., Loukopoulos, T., Khan, S.U., Xu, C.-Z., Zomaya, A.Y.: Survey on grid resource allocation mechanisms. J. Grid Comput., Springer, **12**(2), 399–441 (2014)
2. Sharma, S., Chabbra, A., Sharma, S.: Comparative analysis of scheduling algorithms for grid computing. In: Advances in Computing, Communications and Informatics (ICACCI) Conference, pp. 349–354. IEEE (2015)
3. Kant Soni, V., Sharma, R., Mishra, M.K.: An analysis of various job scheduling strategies in grid computing. In: Signal Processing Systems (ICSPS), 2010 2nd International Conference, vol. 2, pp. 349–354. IEEE (2010)
4. Xhafa, F., Abraham, A.: Computational models and heuristic methods for grid scheduling problems. J. Futue. Gener. Comput. Syst. **26**(4), 608–621. Elsevier (2010)
5. Kolodziej, J.: Evolutionary Hierarchical Multi-Criteria Metaheuristics for Scheduling in Large-Scale Grid Systems. Springer, New York (2012)
6. Babafemi, O., Sanjay, M., Adigun, M.: Towards developing grid-based portals for e-commerce on-demand services on a utility computing platform. J. Procedia **4**(1), 81–87. Elsevier (2013)
7. Desell, T., Newberg, L.A., Magdon-Ismail, M., Szymanski, B.K., Thompson, W.: Finding protein binding sites using volunteer computing grids. In: Gaol, F.L., Nguyen, Q.V. (eds.) 2nd International Congress on Computer Applications and Computational Science, pp. 385–393. Springer, Heidelberg (2012)
8. Sukhija, N., Datta, A.K.: C-grid: enabling iRODS-based grid technology for community health research. In: Information Technology in Bio- and Medical Informatics, pp. 17–31. Springer (2013)

9. Alobaedy, M.M., Ku-Mahamud, K.R.: Scheduling jobs in computational grid using hybrid ACS and GA approach. In: International Conference on Computing, Communications and IT Applications, pp. 223–228. IEEE (2014)
10. Zapfel, G., Braune, R., Bogl, M.: Metaheuristic Search Concepts a Tutorial with Applications to Production and Logistics. Springer, Heidelberg (2010)
11. Xin-She, Yang.: Nature-Inspired Optimization Algorithms. Elsevier, Amsterdam (2014)
12. Tavares Neto, R.F., Godinho Filho, M.: Literature Review Regarding Ant Colony Optimization Applied to Scheduling Problems: Guidelines for Implementation and Directions for Future Research, **26**(1), pp. 150–161, (2013). Elsevier
13. Dorigo, M., Birattari, M.: Ant colony optimization. Encyclopedia of machine learning. Springer (2010)
14. Kang, Q., He, H., Wang, H., Jiang C.: A novel discrete particle swarm optimization algorithm for jobscheduling in grids. In: Natural Computation, 2008. ICNC '08. Fourth International Conference, pp. 401–405 (2008)
15. Imran, M., Hashim, R., Khalid, N.E.A.: An overview of particle swarm optimization variants. Science Direct, pp. 491–496. Elsevier (2013)
16. Ariyasingha, I.D.I.D., Fernando, T.G.I.: Performance analysis of the multi-objective ant colony optimization algorithms for the traveling salesman problem, vol. 23, pp. 11–26. Elsevier (2015)
17. Kousalya, K.: To improve ant algorithm' s grid scheduling using local search. Int. J. Comput. Cogn. 7, 47–57 (2009)
18. Izakian, H., Ladani, B.T., Zamanifar, K., Abraham, A.: A novel particle swarm optimization approach for grid job scheduling. Inf. Syst. Technol. manage. Springer, **31**, 100–109 (2009)
19. Wu, Z., Ni, Z., Gu, L., Liu, X.: A revised discrete particle swarm optimization for cloud workflow scheduling. In: Proceedings of 2010 International conference on Computer Intelligent Security (CIS), pp. 184–188. IEEE (2010)
20. Zhang, L., Chen, Y., Sun, R.: A task scheduling algorithm based on PSO for grid computing. Intern. J. Comput. Intell. Res. **4**, 37–43 (2008)
21. Beegom, A.S.A., Rajasree, M.S.: A particle swarm optimization based pareto optimal task scheduling in cloud computing. In: Advance in Swarm Intelligence Notes on Computer Science, p. 79–86. Springer (2014)
22. Hara, A., Matsushima, S., Ichimura, T., Takahama, T.: Ant colony optimization using exploratory ants for constructing partial solutions. In: Evolutionary Computation (CEC), 2010 IEEE Congress, pp. 1–7. IEEE (2010)
23. Arnaout, J.P., Rabadi, G., Musa, R.: A Two-Stage Ant Colony Optimization Algorithm to Minimize the Makespan on Unrelated Parallel Machines with Sequence-Dependent Setup Times, **21**(6), 693–701. Springer (2010
24. Rais, H.M., Othman, Z.A., Hamdan, A.R.: Improved dynamic ant colony system (DACS) on symmetric traveling salesman problem (TSP). Published in Intelligent and Advanced Systems, ICIAS, pp. 43–48. IEEE
25. Anitha, J., Karpagam, M.: Ant colony optimization using pheromone updating strategy to solve job shop scheduling. In: 7th International Conference on Intelligent Systems and Control (ISCO), pp. 367–372. IEEE (2013)
26. Zhao, N., Wu, Z., Zhao, Y., Quan, T.: Ant colony optimization algorithm with mutation mechanism and its applications. Expert Syst. Appl., Elsevier, **37**(7), 4805–4810 (2010)
27. Ku-Mahamud, K.R.: Ant Colony algorithm for job scheduling in grid computing. In: 2010 Fourth Asia International Conference on Mathematical/Analytical Modelling and Computer Simulation, Sintok, Malaysia, pp. 40–45, 26–28 May 2010
28. Tawfeek, M.A., El-Sisi, A., Keshk, A.E., Torkey, F.A.: Cloud task scheduling based on ant colony optimization. In: 8th International Conference on Computer Engineering and systems, p. 64–69 (2013)
29. Bagherzadeh, J., MadadyarAdeh, M.: An improved ant algorithm for grid scheduling problem using biased initial ants. In: 3rd International Conference on Computer devices, p. 373–378 (2011)

30. Liu, A.L.A., Wang, Z.W.Z.: Grid task scheduling based on adaptive ant colony algorithm. In: International Conference Management e-Commerce e-Government. p. 415–418. IEEE (2008)
31. Laalaoui, Y., Drias, H., Bouridah, A., Ahmed, R.B.: Ant colony system with stagnation avoidance for the scheduling of real-time tasks. In: Computational Intelligence in Scheduling, pp 1–6. IEEE (2009)
32. Xhafa, F., Duran, B., Kolodziej, J.: On exploitation vs exploration of solution space for grid scheduling. In: 3rd International Conference on Intelligent Networking and Collaborative Systems, pp. 164–171. IEEE (2011)
33. Alobaedy, M.M., Ku-Mahamud, K.R.: Scheduling jobs in computational grid using hybrid ACS and GA approach. In: International Conference on Computing, Communications and IT Applications, pp. 223–228. IEEE (2014)
34. Kennedy, J., Eberhart, R.C.: Particle swarm optimization. In: Proceedings of IEEE International Conference on Neural Network, pp. 1948–1995 (1995)
35. Xue, S., Wu, W.: Scheduling workflow in cloud computing basedon hybrid particle swarm algorithm. TELKOMNIKA Indones. J. Electr. Eng. **10**, 1560–1566 (2012)

Survey of Various Real and Non-real-time Scheduling Algorithms in Mobile Ad hoc Networks

Abu Sufian, Anuradha Banerjee and Paramartha Dutta

Abstract Mobile Ad hoc Networks (MANETs) is rapidly gaining popularity due to its infrastructureless and self-configured capabilities. In MANETs, a group of nodes make a network where every node, individually can play sender, receiver or router and the communication is either single hop or multi-hops. So, scheduling is an important part to this kind of network for either route discovery or data packet transfer within this network. All scheduling algorithms in MANETs are either packet scheduling in term of QoS or MAC access. Packet scheduling algorithm is more important than MAC access. Packets are also two types, control packet which are used for route establishment and data packet. Scheduling algorithms can be priority based or non-priority based. We can broadly classify the entire scheduling schemes into two types: Non-real-time scheduling or Real-time scheduling. This paper makes a survey of all kind of scheduling algorithms in Mobile Ad hoc Networks in chronological order with classification of Non-real-time and Real-time scheduling schemes.

Keywords Ad hoc network · Scheduling algorithms · Real-time scheduling · Non-real-time scheduling

A. Sufian (✉)
Department of Computer Science, University of Gour Banga, Malda, WB, India
e-mail: sufian.asu@gmail.com

A. Banerjee
Department of Computer Application, Kalyani Government Engineering College, Kalyani, WB, India
e-mail: anuradha79bn@gmail.com

P. Dutta
Department of Computer & System Science, Visva-Bharati University, Santiniketan, WB, India
e-mail: paramartha.dutta@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_13

1 Introduction

Mobile Ad hoc Network or simply MANETs is a self-configured, infrastructureless, license-free network which are very useful in emergency-like situations, viz., natural disaster, rescue operation, battle field, etc. Here, every node work in Ad hoc basis, that is, every node can play role of sender, receiver or router, and communication can be either single hop or multi hops. The challenge to this kind of network is frequent change of their network topology as every node is dynamic in nature, so route discovery processes require restarting frequently. For that reason burden of every node becomes so high. For this reason, route discovery process as well as packets transfer policy need to be taken carefully with best possible path with best possible way and scheduling algorithm plays an important role to do that. Scheduling algorithms are either packet scheduling in term of QoS or Media Access Control (MAC). Packet scheduling is more critical and vital than MAC access in MANETs. Scheduling algorithm can be priority-based or non-priority-based and it can have preemptive or non-preemptive nature. The choice of scheduling algorithm to determine which queued packet process next which are very significant when traffic load are high. Scheduling algorithm works in between routing agent and MAC sublayer shown in Fig. 1. We can broadly classify all scheduling algorithms in MANETs into Non-real-time and Real-time.

This paper makes a complete survey of several scheduling algorithms in Mobile Ad hoc Network. There are several scheduling algorithms like Non-Priority (FIFO), Priority, Fair, Real-time, Fuzzy logic-based- etc; all are discussed in chronological order with the classification of Non-real-time and Real-time scheduler.

There are so many surveys or reviews carried out on scheduling algorithms in Ad hoc network in the last few years, which are published or discussed on many journal and conferences. Some are: in [1] Byung-Gon Chun and Mary Baker, on their paper did survey on scheduling algorithms particularly based on DSR [2] and GPSR [3] protocol. In [4] Fattah and Leung, categorize the scheduling algorithms as work conserving, where nodes do not sit idle if there is a pending message forwarding task waiting in its message queue and non-work-conserving, when a node might wait expecting higher priority packets even if some messages are still waiting in its message queue. In [5] Enzai et al. presented a survey of various scheduling

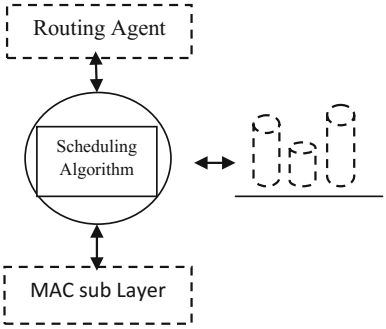


Fig. 1 Position of scheduling algorithm

algorithms and compared their characteristics in terms of advantages and disadvantages. Annadurai [6] classified scheduling algorithms based on whether it is work conserving, priority-based, fair, optimistic and hop dependent. In [7] Garg et al. present a survey of various scheduling techniques based on the parameters of packets they consider and the method of deciding which packet to be served next. Rukmania, Ganesanon [8] makes review on scheduling algorithms that categorizes and prioritizes the real-time traffic with the intention of improving the performance of the real-time applications. They analyzed five types of scheduling algorithms using OPNET [9] simulator. In [10] discussed different types of packet scheduling techniques that are used in wireless sensor networks like FCFS, EDF, Priority, RACE [11], MDP, etc. In [12] author made survey of scheduling algorithms related with short-term, medium-term and long-term scheduler, like FCFS, Round Robin, Priority based, Multilevel queue scheduling, etc. Mohamed Ahmed and W. Jeberson in [13] present an investigation of current state of the art in scheduling and resource allocation mechanisms in ADHOC (MANETT) and WiMAX networks.

We break this article into three sections and the rest of this paper is organized as follows: In Section 2 we make literature survey of present scheduling algorithms and conclusion of this article is in Section 3.

2 Literature Survey of Scheduling Algorithms in MANETS

We made this survey by dividing all scheduling algorithms in mobile Ad hoc networks into Non-real-time and Real-time. In non-real-time, packets are placed in a queue and give services on FCFS basis, short message first or any non-real-time basis. On the other hand real-time basis scheduling, packets are classified into real-time and non-real-time packets, and real-time packet gets served first by using multilevel queue.

2.1 Non-real-time Scheduling

Non-real-time scheduling concentrates several QoS parameter-like Congestion Control, High throughput, minimum end-to-end delay, packet delivery ratio, etc. Several non-real-time scheduling which are currently using or proposed in MANETS are discussed below.

2.1.1 First Cum First Serve (FCFS)

FCFS or FIFO queuing is the simplest scheduling algorithm. There is exactly one message queue in each node. Whenever a new packet arrives at a router, it is placed

at the rear end of the message queue. This scheduling policy take very low computational load. These type are easy to understand and implement as logic behind are very simple. Waiting time can be calculated easily but waiting time is always maximum compare to other non-FCFS schemes. If we want to give different services to different packet in term of times, it would not possible as all the packets placed on single queue.

2.1.2 According to Size and Number of Hops

In this sub sub section, we mentioned different non-real-time scheduling algorithms, where priorities are assigning to the packets depending upon its message size or number of hops to travel. In Smallest Message First (SMF), if a packet belongs to a smaller message, then it will be served before another packet that is part of a larger message. So, after an entire message is fragmented into packets, the total message size must be appended to each of those fragments. Before placing a packet in the queue, the underlying node will compare the size of the parent message with the sizes of those waiting in its queue, to properly place the newly arrived packet in it's queue.

In order to illustrate the SRMF [14] Smallest Remaining Message scheme, we present an example. Say a message m1 is divided into p1 number of packets. Among them i number of packets have been delivered to the destination. So the remaining number of packets is (p1−i). The i+1 th packet arrives at a router where two packets are waiting. The first one belongs to message m2 out of which p2 number of packets have been formed, j no. of packets have been delivered and currently waiting packet is j+1 th. Similarly, the second one belongs to message m3 out of which p3 number of packets have been formed, k no. of packets have been delivered and currently waiting packet is k+1 th. Without any loss of generality, assume that (p1−i) > (p2−j) and (p1−i) < (p3−k). If (p2−j) < (p3−k), then the queue will look like,

$$\begin{array}{ll} (j+1|m2)(i+1|m1)(k+1|m3) & \\ \uparrow & \uparrow \\ \text{Front} & \text{Rear} \end{array}$$

The packet at the front end is served first.

In SHLF scheduling, the packets are prioritized based on the hop count to travel from source to destination. The hop distance from source to destination is attached along with each packet so that the routers can prioritize.

2.1.3 URS and NURS Scheme

In [15] Guo and Kuo proposed two scheduling algorithms URS and NURS. Each node is equipped with a M/M/1 message queue where the router plays the role of a

single server. The packet arrival rate is exponential and the path delay follows normal distribution. The average total delay and re-sequencing delay are calculated based on the distribution of traffic load among the paths. Simulation results show that URS outperforms NURS. URS reduces the re-sequencing delay by increasing the round length.

2.1.4 Fuzzy QoS Controllers Based Priority Scheduler for Mobile Ad hoc Networks

Sun et al. proposed in [16], a priority scheduling scheme in MANETs based on fuzzy QoS controller, where Weighted Round Robin (WRR) packet scheduling was used as the basic scheduling and packets are forwarded based on their weight. Here number of hops on a path, buffer size, queue length, data rate and expiry time determines weight or priority index of a packet using fuzzy logic.

2.1.5 Congestion Aware Scheduling Algorithm (CARE) for MANETs

Z. Chen et.al. proposed Congestion Aware Scheduling Algorithm (CARE) for MANETs in [17]. The authors establish that prioritizing a packet independent of it's neighborhood, does not correctly model a network in real-time, especially when routes are breaking frequently and bandwidth is a rare resource. The load information is selected as congestion indicator which is periodically defined. If the packet arrival rate is higher than packet forwarding rate, then the router is bound to be congested.

They also change the format of RREQ and RREP by adding one extra field that is load information. In their scheme, the control packet gets more priority over data packet and if the priorities are same, tie can be break by FIFO order. When buffer is full, the lowest priority packets are dropped on arrival of higher priority packets to manage the congestion.

2.1.6 Channel Aware Packet Scheduling for MANETs

K.N. Sridhar and Mun Choon Chen have proposed a novel scheduling schemes called channel Aware Scheduling for MANETs (CaSMA) in [18], which incorporate inherent multichip features of MANETs. This algorithm take an unique approach "Channel-Aware" here channel aware means knowledge of current status of channel which refers to quality of channel in terms of suitable metrics.

CaSMA focused on end-to-end channel quality as a function of route's lifetime. This will avoid congestion, gives maximum speed and reduce packet drop rates.

2.1.7 ELQS: An Energy Efficient and Load Balanced Queue Scheduling for MANETs

Yin. Y and Yang. X have proposed ELQS scheme in [19], especially to reduce energy consumption as well as wastage, as much as possible, since the nodes in Ad hoc networks are battery powered and hence energy constrained. A packet travelling through a route containing low energy nodes are given high priority.

2.1.8 Dynamic Congestion Control with Multihop Scheduling Algorithm for MANETs

Naveen. G and HemalBabu. H in [20] proposed scheduling scheme in MANETs, where congestion control, throughput and end-end delay are considered. They make this algorithm distributed to different sub algorithms, easy to implement and use stochastic dominance method to analyze the end-to-end delay.

They use different sub-scheduling for different QoS: Rate-Based Scheduling algorithm for time slot allocation. Window-Based Flow Control Algorithm for congestion control where node must need to acknowledge to source node each time, to bound number of packets, so buffer overflow can be controlled. Routing Model for data transmission such as Channel Reservation to reserve the channel over wireless medium, Connection Management and Maintenance of Route Reservation, and Route Failure Handling. The integrated components work together to greatly reduce end-to-end delay and efficiently utilize bandwidth in wireless network.

2.1.9 Fair Scheduling

In this sub sub section, we deal with different Fair scheduling algorithms [21–24]. Much research effort has been made on research in scheduling focused on fairness issues. Basically two types of FAIR scheduling are there: one is timestamp based and other is credit based. In timestamp based mechanism packet is assigned a start tag and a finish tag. Anyone of them may be used for forwarding. For example, the oldest packet, i.e., the one with smallest start tag is served first.

On the other hand Maxmin Fair scheduling in [25], describes that every session always has a packet to transmit. Each node gets service token in a session in round robin fashion. A token is removed from both source and destination when they get service.

2.1.10 Dev-Ns

A. Banerjee and P. Dutta in [26] proposed scheduling scheme, DEV-NS: Delay-Efficient Energy and Velocity Conscious non-preemptive Scheduler, that instructs each node to assign weight to its uplink neighbors at regular intervals.

Message forwarding request from the uplink neighbor with highest weight is served first. In DEV-NS if node N_i sends some message to node N_j which N_j is supposed forward to node N_k in a communication route then N_i can send the next packet to N_j only after N_k acknowledge of earlier packet to N_i. The key ideas are: first identifies vulnerable nodes by some deciding factors such as residual energy, relative velocity, and number of alive communication session in certain period of time, then calculate vulnerability index and delay index. By using these parameters select best scheduler.

Purpose of DEV-NS is to serve those packets which are expected to suffer from minimum delay. Link breakage may take place due to complete exhaustion of battery power of a node or node goes out of radio range. This scheme is concerned about cost of route discovery. This scheme not only produces a delay-efficient schedule but also decreases cost and improved throughput. The simulation results of this scheme also claim the same.

2.2 Real-Time Scheduling

Real-time scheduling in MANETs, divides all packets into real-time packets and non-real-time packets. Real-time packets are time bounded so it needs to get service within time bound. Almost all real-time scheduling are using or proposed so far, are discussed below:

2.2.1 Real-Time Traffic Support in Large-Scale MANETs

Yuan Sun et al. proposed a real-time traffic support scheme in [27]. The call setup protocol in the IP layer distinguishes real-time flow from non-real-time ones. The underlying MAC layer assigns more priority to real-time ones. As soon as a forwarding request arrives at a router, the call set up process build a path to the destination. For real-time packets, the flow admission control component tries to fulfill the forwarding request without any degradation of the present flow of packets. This requires estimation of resources in the routers connecting the source with destination.

2.2.2 Opportunistic Packet Scheduling

Many research works has been proposed on opportunistic packet scheduling [28–30]. Assume that nodes 0, 1, 2, 3 and 4 are there. Excluding node 0, the queue of other nodes contains nodes 1, 2, 3 and 4 in that order (without any loss of generality). Node 0 transmits RTS or request to send. Among nodes 1, 2, 3 and 4, node 1 is placed at the front end of the queue and it will send CTS. After the current

session, node 1 will be placed at the rear end and the order will look like 2, 3, 4 and 1. Next time, node 2 will have highest priority.

The concept of cooperation among neighbor nodes is introduced by distributed cooperative and opportunistic scheduling (COS). For consecutive data transmission of the unscheduled transmitters, extra intervals are inserted so that the highest priority transmitter gets access to the wireless channel. On the other hand in distributed opportunistic scheduling (DOS), many transmitters compete to gain access to a wireless channel and anyone gets it randomly. Opportunistic Scheduling for Ad hoc Network (ROSA) in [31], initially it was designed for infrastructure network, later adopted Ad hoc networks. This adopted scheme work on error propagation as well as evaluating the link fitness and fairness of the scheduling policy. Robustness of the network is improved by imposing an upper limit on the traffic load at MAC sub layer.

2.2.3 Cross-Layer Packet Scheduling

K.J. An and H. Song in [32], has proposed a Cross Layer Packet Scheduling and routing algorithm that effectively delivers delay-sensitive data in MANETs by using multiple paths. Here, sum metrics like packet urgency, node urgency and route defined and effective algorithm are designed by using these metrics. The packet urgency metric are defined first according to end-to-end delay and network topology. This cross-layer scheme tries to spread traffic load over entire network. This can significantly improve end-to-end delivery in the specified delay.

2.2.4 Edf

Earliest Deadline First (EDF), as the name suggests, each packet is assigned a deadline and highest priority is assigned to the packet having the smallest deadline. Deadline specifies the upper limit of time by which a packet must reach its destination. Y. Debhi and N. Mikou in [33], proposed another heuristic EDF called "Enhanced EDF". It modifies EDF to enhance QoS for important classes without degradation of less important packet queue by using sigmoid priority function. It shows by simulation that this version of EDF performs better in nearly overloaded situation.

2.2.5 Dynamic Multilevel Priority (DMP)

Nidal Nasser et al. in [34] proposed a scheduling scheme based on dynamic multilevel priority (DMP). In DMP scheduling, the nodes are in hierarchical format where three level of queue are used. The leaf nodes should have two queues for both real-time and non-real-time packets, so they do not receive data from another node. The nodes which equally distance from base node are at same level. It also

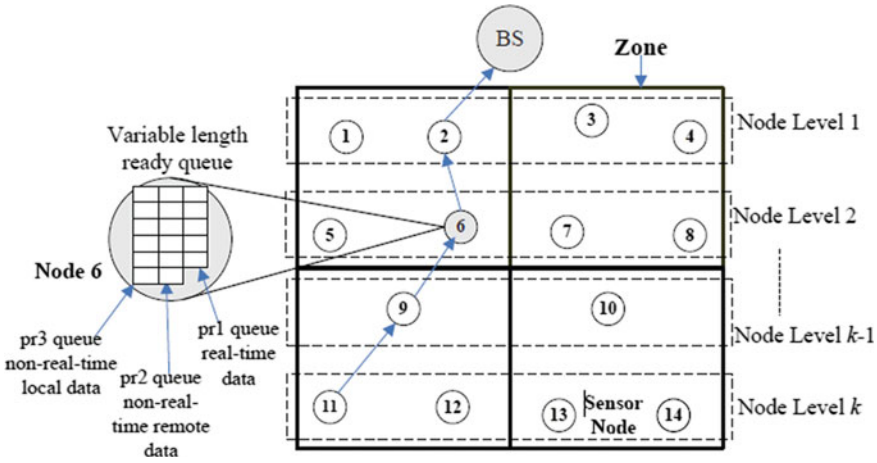


Fig. 2 DMP packet scheduling [34]

used TDMA for those nodes are situated at different levels. Intermediate nodes have more task than leaf as they are connecting path.

In this Fig. 2 of DMP, three priority queues p1, p2, and p3 are used. In this figure for node6, the real-time packet goes to p1 which is highest priority. The non-real-time data packet from lower nodes such as from node9 or node11 goes to p2 and local non-real-time data packet goes to p3. We can see from the figure that entire networks break into different zones. This figure also shown node1, node2, node3 and node4 are equal distances from base node (BS), so they are in same level level 1, similarly level 2, level 3, etc., contains those nodes which are equally distances from base node (BS) respectively.

2.2.6 Low Latency Queuing (LLQ) Algorithm

In [35] LLQ scheme based on fair queuing scheduling delay-sensitive traffic is given special handling over other traffic. The LLQ uses a single priority queue within which packets belonging to different priority classes are embedded. For example, voice data is assigned the highest priority, then videos and then other data packets. Actually the packet with highest delay sensitivity is forwarded first. The key difference between LLQ and PQ is that in PQ, the low priority queues starve which does not happen in LLQ (Fig. 3).

2.2.7 Delay Sensitive Packet Scheduling Algorithm for MANETs

In [36] Vaidhagi et al. proposed a delay-sensitive packet scheduling algorithm in MANETs to deliver the delay sensitive data without any delay. In order to achieve

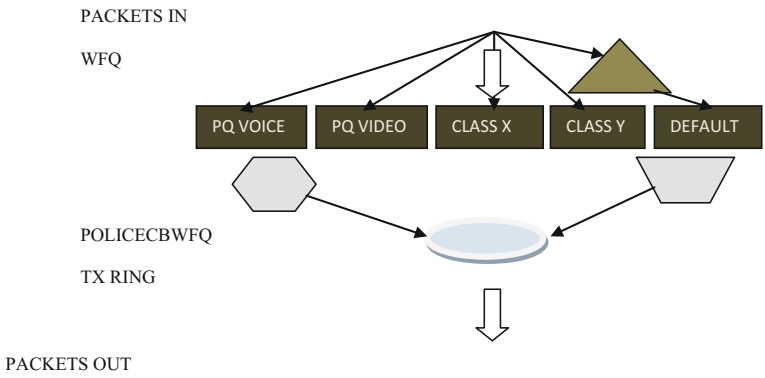


Fig. 3 LLQ

this they used some urgency metrics such as: Packet Urgency, Node Urgency, and Route Urgency.

These metric needs to combine after that this combination metric determine the order of transmission. Here little modification of AODV [37] protocol done by including RREQ packet header to construct route. The route selection is done by minimum packets drop rate. A threshold value of dropped packets is predetermined based on the network parameters. If in a particular route, the minimum number of packet drop is less than threshold, then the route is considered stable for delay-constrained communication.

2.2.8 FSRP

A. Banerjee et al. on their article in [38] proposed one scheduling scheme called FSRP that will prioritize the real-time message packet based on deadline to packets delivery, relative velocity between consecutive routers in the path along with their residual energy.

According to this scheme, as soon as real-time message packet generated by source N_s for destination N_d arrives at router N_i at time t, N_i first checks whether it is really possible to forward the data packet to the destination N_d within the deadline. If it is impossible then the packet is drooped by N_i, because if the time and energy will be wasted in the route discovery process to build a new path from that router to the destination.

To define the stable path the authors used Real-Time Fuzzy controller (RTF) by using Deadline Impact (LI), Stability Impact (SI) and Source Priority (SP) and these are formulated by fuzzy rule. Then by combining these parameters get RTF. RTF gives rank to all possible paths, and then according to this rank path are selected for real-time traffic. Simulation result in NS-2 [39] also shown improvement in this article where FSRP embedded with AODV [37] and FAIR [40] and compared with other scheduler embedded with same protocols.

3 Conclusion

In this article we have discussed real-time and non-real-time scheduling algorithms in Ad hoc Networks. We have also seen that non-real-time scheduling algorithm's QoS in terms of Congestion Control, Reliability, Minimum end-to-end delay, Packets deliver ratio, etc., whereas real-time scheduling concentrate on real-time data transmission without degrading other QoS. As real-time applications rapidly growing in modern day arena, so in order to support real-time applications, scheduling algorithm needs much improvement without affecting other QoS in Ad hoc Networks. This article might help to do future research on scheduling algorithm in Ad hoc Networks.

References

1. Chun, B., Baker, M.: Evaluation of packet scheduling algorithms in mobile ad-hoc networks. ACM SIGMOBILE Mob. Comput. Commun. Rev. 36–49 (2002)
2. Johnson, D.B., Maltz, D.A., Hu, Y.C., Jetcheva, J.G.: The dynamic source routing protocol for mobile ad hoc networks. IETF Internet Draft, Mobile Ad-hoc Network Working Group, IETF (2002)
3. Karp, B., Kung, H.T.: GPSR: greedy perimeter stateless routing for wireless networks. In: Proceedings of 48 Mobile Computing and Communications Review, vol. 6, no. 3 Boston, MA (2000). ACM MOBICOM
4. Fattah, H., Leung, C.: An overview of scheduling algorithms in wireless multimedia networks. IEEE Wireless Commun. J. **9**, 76–83 (2002)
5. Enzai, N.I.M., Rais, S.S., Darus, R.: An overview of scheduling algorithms in mobile ad-hoc networks. In: International Conference on Computer Applications and Industrial Electronics (ICCAIE 2010), Kuala Lumpur, Malaysia, 5–7 Dec 2010
6. Annadurai, C.: Review of packet scheduling algorithms in mobile ad hoc networks. Int. J. Comput. Appl. **15**(1) (2011)
7. Garg, K., Pal, R.: Scheduling algorithms in mobile ad hoc networks. Int. J. Comput. Sci. Appl. (TIJCSA). **1**(5) (2012)
8. Rukmania, P., Ganesan, R.: Scheduling algorithm for real time applications in mobile ad hoc network with opnet modeler. In: International Conference on Design and Manufacturing, IConDM (2013)
9. http://www.sce.carleton.ca/faculty/lambadaris/courses/5001/opnet_tutorial.pdf
10. Sangvikar Praddyumna, P., Sandip, K.: Review paper on packet scheduling in wireless sensor network. Int. J. Emerg Tech. Adv. Eng. **4**(10) (2014)
11. Mizanian, K., Hajisheykhi, R., Baharloo, M., Jahangir, A.H.: RACE: a real-time scheduling policy and communication architecture for largescale wireless sensor networks. In: Proceedings of Communication Network Services Research Conference, pp. 458–460 (2009)
12. Tharani, K., Kumuthini, C.: A survey on various scheduling techniques based on Qos Factors in Manet. Int. J. Sci. Res. **3**(11) (2014)
13. Ahmed, M., Jeberson, W.: Review on scheduling and resource allocation management techniques in ad-hoc and wimax networks. Int. J. Emerg. Trends Technol. Comput. Sci. (IJETTCS) **4**(6) (2015)
14. Modiano, E.: Scheduling packet transmissions in a multi-hop packet switched network based on message length. In: Proceedings of the Sixth International Conference on Computer Communications and Networks, pp. 350–357 (1997)

15. Guo, Y.F., Kuo, G.S.: A packet scheduling framework for multipath routing in mobile ad-hoc networks. In: IEEE Wireless Communications and Networking Conference, pp. 1081–1086 (2004)
16. Sun, B., Li, L., Gui, C.: Fuzzy QoS controllers based priority scheduler for mobile ad hoc networks. In: International Conference on Mobile Technology, Applications and Systems, p. 5 (2005)
17. Chen, Z., Ge, Z., Zhao, M.: Congestion aware scheduling algorithm for MANET, WiCOM, pp. 1–5 (2006)
18. Sridhar, K.N., Chan, M.C.: Channel-aware packet scheduling for MANETs. In: World of Wireless, Mobile and Multimedia Networks, pp. 1– 9 (2008)
19. Yin, J., Yang, X.: ELQS: An energy efficient and load balanced queue scheduling algorithm for mobile ad-hoc networks. CMC **2**, 121–126 (2009)
20. Naveen, G., HemalBabu, H.: Dynamic congestion control with multihop scheduling algorithm for mobile ad-hoc network. Int. J. Innovative Res. Comput. Commun. Eng. **2**(1) (2014)
21. Lu, S., Bharghavan, V., Srikant, R.: Fair scheduling in wireless packet networks. In: Proceedings of ACM SIGCOMM, Cannes, France (1997)
22. Nantagopal, T., Kim, T., Gao, X., Bharghavan, V.: Achieving MAC layer fairness in wireless packet networks. In: Proceedings of ACM MOBICOM, Boston, MA (2000)
23. Luo, H., Lu, S., Bharghavan, V.: A new model for packet scheduling in multihop wireless networks. In: Proceedings of ACM MOBICOM, Boston, MA (2000)
24. Semeria, C.: Supporting differentiated service classes: queue scheduling disciplines. In: Juniper Networks, pp. 11–14 (2001)
25. Tassiulas, L., Sarkar, S.: Maxmin fair scheduling in wireless ad hoc networks. IEEE J. Selected Areas Commun. **23**(1) (2005)
26. Banerjee, A., Dutta, P.: Delay-efficient energy and velocity conscious non-preemptive scheduler (DEV-NS) for mobile ad hoc networks. Int. J. Adv. Netw. Appl.(**5**)4 (2014)
27. http://broadnets.org/2004/workshop-papers/Broadwim/broadwim2004-Paper13-sun.pdf
28. Wang, J., Zhai, H., Fang, Y.: Opportunistic packet scheduling and media access control for wireless LANs and multi-hop ad hoc networks. In: IEEE Commun. Soc. WCNC (2004)
29. Chen, Q., Zhang, Q., Niu, N.: Opportunistic link scheduling with qos requirements in wireless adhoc networks. In: Proceedings of IEEE Communications Society Subject Matter Experts for Publication in the ICC (2007)
30. Zheng, D., Ge, W., Zhang, J.: Distributed opportunistic scheduling for ad-hoc communications: an optimal stopping approach. MobiHoc'07, ACM (2007)
31. Sun, Y., Li, V.O.K., Leung, K-C.: Distributed Opportunistic Scheduling in Multihop wireless Ad Hoc Networks, IEEE (2008)
32. An, K.J., Song, H.: An effective cross-layer packet scheduling and routing algorithm for delay-sensitive media transmission over MANET. In: IEEE International Conference on Communications, ICC '09, (2009) and Proceedings of India Annual Conference, IEEE INDICON, pp. 386–391 (2004)
33. Dehbi, Y., Mikou, N.: Priority assignment for multimedia packet scheduling in MANET. In: International Conference on Signal Image Technology and Internet Based Systems, SITIS '08, pp. 32–37 (2008)
34. Nasser, N., Karim, L., Taleb, T.: Dynamic multilevel priority packet scheduling scheme for wireless sensor network. IEEE Transact. Wireless Commun. **12**(4) (2013)
35. LowLatency Queuing Algorithm (LLQ), http://www.cisco.com/en/US/docs/ios/12_0t/12_0t7/feature/guide/pqcbwfq.pdf
36. Vaidhegi, E., Padmavathy, C., Priyanga, T., Priyadharshini, A.: Delay sensitive packet scheduling algorithm for MANETs by cross layer. Int. J. Innovative Res. Adv. Eng. (IJIRAE) **1**(1) (2014)
37. Perkins, C.E., Royer, E.M.: Ad hoc on demand distance vector routing, mobile computing systems and applications. In: Proceedings of Second IEEE Workshop on WMCSA '99, pp. 90–100 (1999)

38. Banerjee, A., Dutta, P., Sufian, A.: Fuzzy-controlled scheduling of real time data packets (FSRP) in mobile ad hoc networks. Int. J. Comput. Sci. Mob. Comput. **5**(5), 507–513 (2016)
39. http://www.isi.edu/nsnam/ns/tutorial
40. Banerjee, A., Dutta, P.: Fuzzy-controlled adaptive and intelligent route selection (FAIR) in ad hoc networks. Euro. J. Sci. Res. **45**(3), 367–382 (2010)
41. http://www.tutorialspoint.com/operating_system/os_process_scheduling_algorithms.htm

Double Ended Bottom-Up Approach of Data Warehouse Architecture Employing Country ISO Code Engendered Cryptographic Algorithm

Rajdeep Chowdhury, Ipshita Bhattacharya, Nirmita De and Subhajit Saha

Abstract Data Warehouse is an incorporated database premeditated to expand while crafting decision and crisis resolving, espousing exceedingly condensed fact. The amalgamation of security modus operandi with data warehouse is a unanimously established premise for contemporary researchers. The crucial tip of the projected effort through the formulated work emphasizes on the design and implementation of the novel Double Ended Bottom-Up Approach architecture and the integration of the cryptographic modus operandi amid it having precise intent of augmenting safety and performance therein. The proposed cryptographic algorithm has been engendered based on the concept of employing Country ISO Code within. Devising the inventive algorithm ensures considerable diminish in admittance time, considering apt retrieval of the entire indispensable information. Related appliance of the proposed work is in a diversity of organization, where accrual of plausible record is of extreme insinuation. The collection of organization embraces an amalgamation of scholastic organizations, business houses, curative medicinal enterprises, defense and security agencies with sensitive data, classified enterprises and forwards.

Keywords Data warehouse · Encryption · Country ISO code · Modified ASCII · Transposition · Double ended bottom–up approach

R. Chowdhury (✉)
Department of Computer Application, JIS College of Engineering, Block-A, Phase-III, Kalyani, Nadia 741235, West Bengal, India
e-mail: dujon18@yahoo.co.in

I. Bhattacharya · N. De · S. Saha
Department of Electronics and Communication Engineering, JIS College of Engineering, Block-A, Phase-III, Kalyani, Nadia 741235, West Bengal, India

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_14

1 Introduction

In this paper, both encryption and decryption modus operandi have been implemented and employed for performance enhancement of the data warehouse architecture [1–10]. Data warehouse execution augmentation has been allied amid employment of the engendered cryptographic technique amid its structure [11–15]. In the devised cryptographic algorithm, the ASCII Table has been modified by adding 270 to apiece original ASCII Table value.

The key is engendered employing an iterative formula and a series of numbers are engendered there on. The engendered key value is a combination of decimal and whole numbers. For simplicity, the decimal numbers are avoided and whole numbers are employed. The number engendered is the blend of two digits. A Country ISO Code is incorporated in the code. Using the table, the encrypted value is obtained. Hence, the first digit is taken as row value and the second digit is taken as column value. With the aid of row value and column value, a unique ISO Code is engendered. Apiece character of the password entered is replaced by the unique ISO Code. Again, the ISO Code Table is modified into a new table consisting of some special characters. These characters are being arranged in a zigzag pattern employing the Transposition method and the resultant value is the cipher text [14, 15].

During decryption modus operandi, after engendering the cipher text, the entire process is reversed employing the identical modus operandi [14, 15].

The text is split in a particular manner to perform the zigzag Transposition method. Then, the text is interpreted in the Extended ASCII Table. Apiece character of Extended ASCII Table represents one of the Country ISO Code.

After obtaining the ISO Code, the next step is to locate the row value and the column value of apiece ISO Code from the given table. The first digit is the row number and the second digit is the column number of the obtained value. The value obtained is put in the equation where the key is engendered, but in the reverse manner. Next, the ASCII value of the character is retrieved and thus retrieval of original text is ensured.

The employment of the Country ISO Code engendered cryptographic algorithm amid the Double Ended Bottom-Up Approach Data Warehouse Architecture ensures significant performance enhancement [1–8].

2 Literature Review

The segment accentuate on interrelated effort obtainable in the midst of identical genus, abet in devising the said manuscript [1–12]. Fairly correlated workings on data warehouse and its safety trial are existent and have formerly been accomplished, although the specific and apposite realization of the novel proposed algorithm is exceedingly pleasing for contemporary researchers. In [1] it has been

observed that the work is based on Pseudo Mesh Schema Architecture signifying that instead of a Centralized Data Warehouse, all the data would be residing over geographically distant Work Station(s) or Site(s). In [2] the decisive lean while devising the work is incorporated on endowing with an improved depiction with the intent of better safety trial, ascertaining superior ease of access employing Hashing method. In [8] it is identified that all the past work that has been influenced in data warehouse endow to safety concern, thereby ensuring the need and the implementation. The work deals with presently obtainable data security techniques focusing on precise issues, concerning their exercise in data warehousing environment. In [10] it has been observed that there have been various alterations in security modus operandi of Data Warehouse, along with alterations in strategies for development. These are few papers that have been studied as part of Literature Review during the conceptualization of the formulated paper and these studied papers played a noteworthy role in finalization of the proposed work.

3 Proposed Work

To ensure convenient identification of the entire proposed work, a flow chart has been designed to elucidate (Fig. 1).

The focal point of the cryptographic algorithm is the Country ISO Code—which is in the form of a 10 × 10 matrix. It is incorporated in the Data Warehouse Architecture in the form of a slide diverse security mechanism.

As the Data Vault and the Data Mart are quite a system with confidentiality, it needs a burly security modus operandi and righteous authentication mechanism. Hence, an innovative algorithm is devised to avoid the unwanted authentication and information hacking by malicious intruders. The integrative blend of Country ISO Code and a splinter of Computer Science and Engineering, particularly cryptographic algorithm is an inventive progression.

In the subsequent splinter, the projected effort has been precisely indicated and the initiation equips the formulation of the proposed innovative Country ISO Code engendered cryptographic modus operandi, as avowed beneath for expediency.

3.1 Salient Points of Country ISO Code Engendered Cryptographic Technique

1. The projected modus operandi is modeled on Country ISO Code.
2. The Country ISO Code Table has been employed for reference in encryption modus operandi.
3. The entire table is in the form of 10 × 10 matrix. By employing the table, the encrypted value is obtained with the aid of a special key formulation.

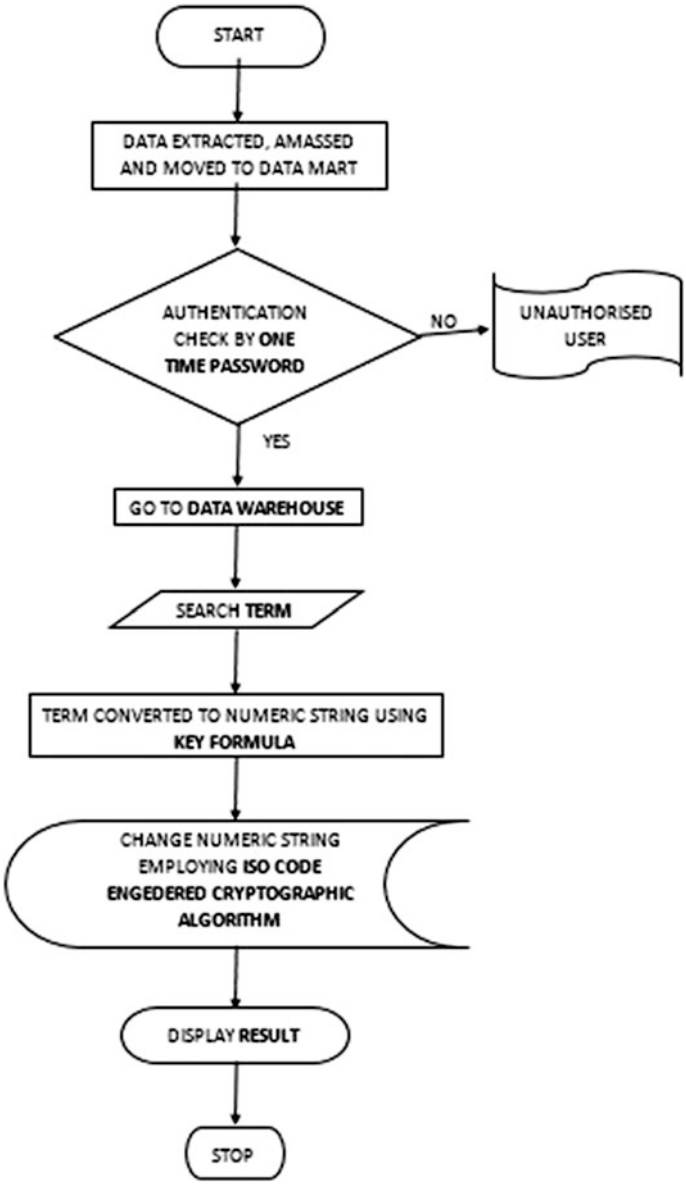


Fig. 1 Flowchart of the proposed work

4. Apiece ISO Code of the matrix is replaced by a special character.
5. The obtained code is being arranged in a zigzag pattern to perform Transposition method and the resultant value is the cipher text.
6. The numeral characters present in plain text is equivalent to numeral characters in cipher text.

3.2 ASCII Table Modification

The ASCII Table has been customized and integrated throughout encryption modus operandi as depicted beneath (Fig. 2).

The Modified ASCII Table exemplifies an effortlessly constructed modus operandi employed to amend the table,

$$A = a + 270 \tag{1}$$

where,

A Latest decimal spot of the character
a Previous decimal spot of the character

Consequently, Modified ASCII Table is created wherein amendments of decimal worth of the characters have been ensured.

DEC	HTML	DEC	HTML	DEC	HTML	DEC	HTML	DEC	HTML	DEC	HTML	DEC	HTML	DEC	HTML
270	NULL	286	DLE	302	SPIC	318	0	334	@	350	P	366	‘	382	p
271	SOH	287	DC1	303	!	319	1	335	A	351	Q	364	a	383	q
272	STX	288	DC2	304	“	320	2	336	B	352	R	368	b	384	r
273	ETX	289	DC3	305	#	321	3	337	C	353	S	369	c	385	s
274	EOT	290	DC4	306	$	322	4	338	D	354	T	370	d	386	t
275	ENQ	291	NAK	307	%	323	5	339	E	355	U	371	e	387	u
276	ACK	292	SYN	308	&	324	6	340	F	356	V	372	f	388	v
277	BEL	293	ETB	309	SPACE	325	7	341	G	357	W	373	g	389	w
278	BS	294	CAN	310	(	326	8	342	H	358	X	374	h	390	x
279	HT	295	EM	311	)	327	9	343	I	359	Y	375	i	391	y
280	LF	296	SUB	312	*	328	:	344	J	360	Z	376	j	392	z
281	VT	297	ESC	313	+	329	;	345	K	361	[	377	k	393	{
282	FF	298	FS	314	,	330	<	346	L	362	\	378	l	394	|
283	CR	299	GS	315	-	331	=	347	M	363	]	379	m	395	}
284	SO	300	RS	316	.	332	>	348	N	364	^	380	n	396	~
285	SI	301	US	317	/	333	?	349	O	365	_	381	o	397	DEL

Fig. 2 Modified ASCII Table

3.3 Encryption Modus Operandi

Crafting of encryption key has been ensured employing an effortless modus operandi;

$$K = 100 * \cos(A) \tag{2}$$

where,

K A number signifying the value to be incorporated in the ISO Code table
A Latest decimal spot of the character

Observances indicate that employing the formula implies that the worth of K always turns out to be decimal in the manner of ab.cd….

Here, only the integer part of the obtained value, discarding the fractional part is considered. Moreover, the negative magnitude is not considered, instead the mod value is considered. Therefore, the first digit 'a' implies the apparent row numeral to be considered and the second digit 'b' implies the apparent column value. From these two digits, the specific ISO Code is obtained (Figs. 3 and 4).

3.4 Sequential Attainment of Cipher Text

1. Primarily, particular character allusion and its decimal worth in accord to the Modified ASCII table have been executed.
2. Formula (1) has been employed to attain a decimal worth as expressed formerly. Allusion to initial numeral acquired to the precise row and second numeral to

	0	1	2	3	4	5	6	7	8	9
0	AF	DZ	BE	BW	IO	VG	BN	CM	AL	TD
1	BF	CA	KY	CC	DJ	HR	KM	BY	AS	CN
2	CO	CX	DK	ED	AW	AZ	BB	KH	TL	DM
3	EC	SV	CD	GQ	CW	BI	AG	AO	ZW	PH
4	WF	VN	UY	FK	EE	GU	HT	IN	IM	CI
5	JE	KZ	CI	XN	LA	MW	YT	AQ	BS	AZ
6	NZ	KP	PK	CG	RO	KN	PM	MF	VC	ST
7	CY	AU	PS	MP	RE	LC	VU	SM	KR	TJ
8	AE	VL	ZM	YE	VA	HU	IE	AQ	BT	AH
9	KG	LV	XK	JO	FR	MM	AN	NL	OM	BL

Fig. 3 Country ISO Code Table

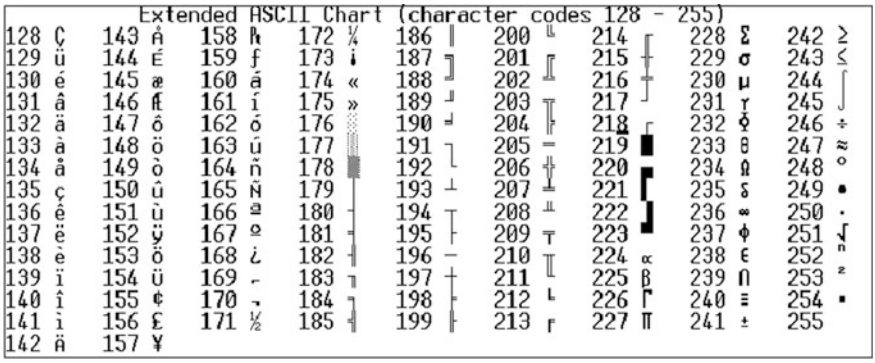


Fig. 4 Extended ASCII Table

the precise column ensures the exact position of the 10 × 10 matrix in the ISO Code table, which is the fundamental code that is requisite for employment.

3. The fundamental code consisting of two alphabets acquired from the preceding table is ascertained.
4. After acquiring the fundamental code, it is replaced with the special characters.
5. The special characters are separated in a specific manner by employing Transposition method using zigzag modus operandi. Thus, couple of separate sequence of special characters is obtained.
6. The couple of sequences are concatenated and finally the cipher text is obtained.

3.5 Decryption Modus Operandi

Decryption is the method of converting from code to plain text. For decryption modus operandi, the engendered cipher text is split in a manner to perform zigzag Transposition method. The obtained result is compared with the Extended ASCII table, wherein apiece character of the Extended ASCII table represents a country ISO code. Thus, the row and the column value of apiece ISO code is obtained. The first digit is considered as the row numeral and the second as the column numeral. The value is placed amid the key equation but in the reverse manner and thereby ensures retrieval of the original text.

Example Demonstration—PROFESSOR@15

See Figs. 5 and 6.

PLAIN TEXT	MODIFIED ASCII TABLE	KEY VALUE	ISO CODE	SPECIAL CHARACTER REPRESENTATION
P	350	28	TL	r
R	352	98	OM	ϴ
O	349	91	LV	N
F	340	75	LC	O
E	339	93	MM	ć
S	353	41	VN	#
S	352	41	VN	#
O	349	91	LV	N
R	352	98	OM	ϴ
@	334	54	LA	ì
1	319	12	KY	*
5	323	83	YE	à

Fig. 5 Table for encryption modus operandi

CIPHER TEXT	SPECIAL CHARACTER REPRESENTATION	COUNTRY ISO CODE	KEY VALUE	MODIFIED ASCII TABLE	PLAIN TEXT
r	r	TL	28	350	P
ō	ϴ	OM	98	352	R
c	N	LV	91	349	O
N	ō	LC	75	340	F
ϴ	c	MM	93	339	E
à	#	VN	41	353	S
ϴ	#	VN	41	353	S
N	N	LV	91	349	O
#	ϴ	OM	98	352	R
#	ì	LA	54	334	@
ì	*	KY	12	319	1
*	à	YE	83	323	5

Fig. 6 Table for decryption modus operandi

4 Result Analysis

For utmost ease in understanding, the entire proposed cryptographic algorithm have been designed and implemented by scripting through 'C' program and a real time example analysis via screenshot have been detailed below in this section, for ease in reference (Fig. 7).

Also in this section, the Encryption time and the Decryption time for.doc files have been compared for ICCECA (ISO Country Code Engendered Cryptographic Algorithm) with respect to AES and Triple DES. The tabular contrast and the graphical exemplification have been ascertained (Table 1 and Fig. 8).

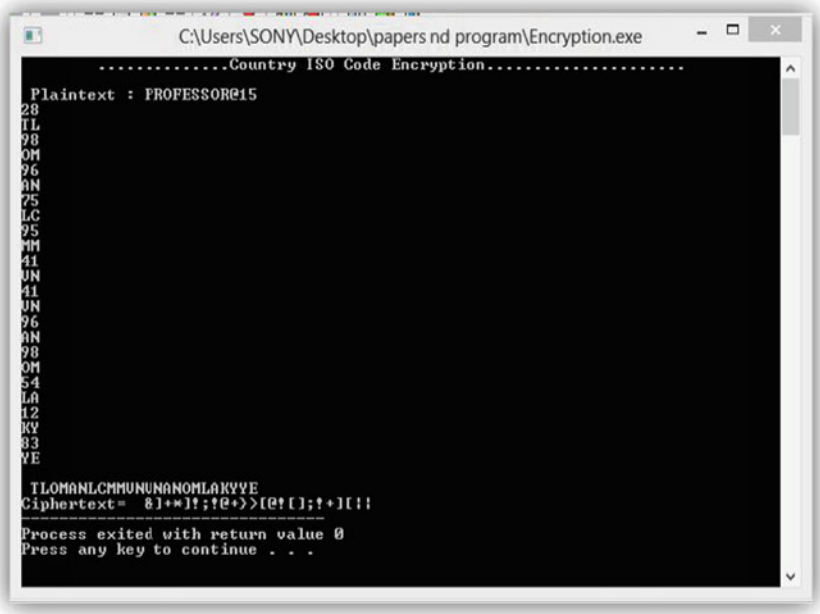


Fig. 7 Screenshot of program output

Table 1 Encryption time and decryption time for .doc files

Sl.	Source File Name	Source File Size	AES (In m. sec.)		TDES (In m. sec.)		ICCECA (In m. sec.)	
		(In Bytes)	Encryption	Decryption	Encryption	Decryption	Encryption	Decryption
1	File1.doc	25,065	16	0	15	0	15	0
2	File2.doc	29,525	17	14	16	15	16	22
2	File3.doc	50,688	22	2	16	0	21	47
4	File4.doc	62,192	17	21	15	0	21	21
5	File5.doc	126,660	25	21	15	16	47	22
6	File6.doc	251,222	19	22	15	22	78	78
7	File7.doc	555,256	21	22	22	15	78	62
8	File8.doc	659,568	74	77	21	21	102	125
9	File9.doc	1,120,596	199	122	99	21	125	156
10	File10.doc	1,595,905	221	261	99	112	196	209
11	File11.doc	2,019,228	204	217	122	125	262	226
12	File12.doc	2,286,105	222	247	140	146	229	294
12	File13.doc	2,625,252	424	518	199	220	259	216
14	File14.doc	2,052,850	448	466	199	201	262	295
15	File15.doc	2,256,556	498	480	214	212	229	295
16	File16.doc	2,655,668	551	512	242	250	292	462
17	File17.doc	5,066,056	621	541	266	249	242	259
18	File18.doc	5,562,856	687	692	202	291	259	416
19	File19.doc	5,655,556	956	924	229	221	296	292
20	File20.doc	5,256,605	997	927	222	216	422	446

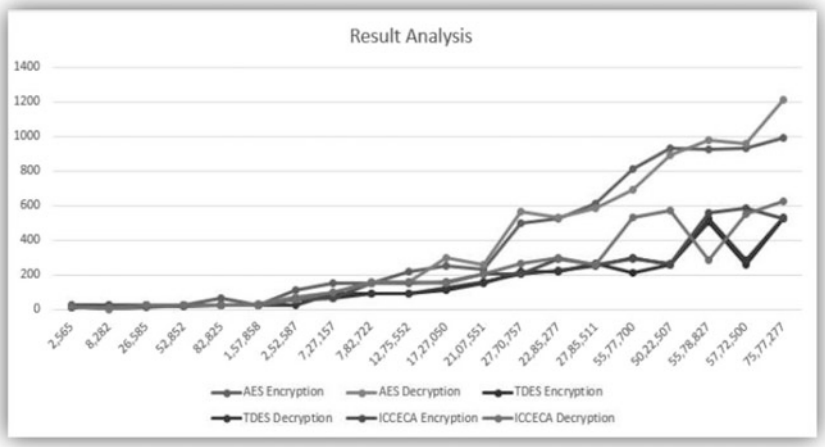


Fig. 8 Result analysis

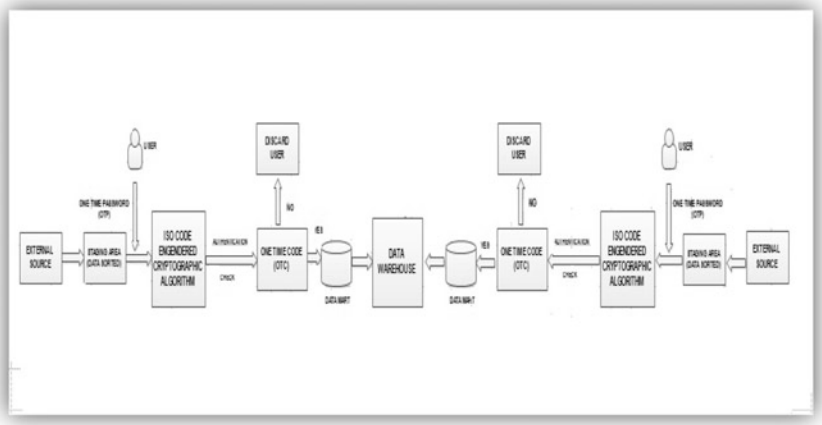


Fig. 9 Working of double ended Bottom-Up approach of data warehouse architecture

5 Working of Double Ended Bottom-Up Approach of Data Warehouse Architecture Employing Country ISO Code Engendered Cryptographic Algorithm

To ensure convenient identification, Fig. 9 elucidates requisite working of the data warehouse architecture.

Step by step elucidation of the working of Double Ended Bottom-Up Approach of Data Warehouse Architecture Employing Country ISO Engendered Cryptographic Algorithm is stated for ease in reference and understanding, namely:

Step 1: Initially, data and request do come from a variety of External Sources placed in both faces of the structure, performing as Double-Sided Queue.
Step 2: Data is amassed in the Staging Area and categorized accordingly in both faces.
Step 3: Then the data is loaded into the Data Marts. At this stage, there is a security check point. Before entering the Data Mart, user has to furnish a Password at both ends and that Password is encrypted and checked via Encryption Algorithm. After that, a One Time Code is sent to the user which determines the authenticity of the user. If the code is not sent, then that user is considered as unauthorized and is not permitted to proceed further.
Step 4: When the authenticity is ensured, the data is entered into the Data Marts from the Staging Area, whose content must be updated and monitered on regular basis.
Step 5: After entering the dedicated Data Marts, information are loaded into the Data Warehouse, where enormous amount of data and information are processed all the time. The structure is termed as Bottom Up Approach as the flow is from the smaller segments to the larger segments and there exists a special approach where the Double Sided Sequence is introduced for ease of processing.
Step 6: In this model, as data do come from both faces, the collision of enormous quantity of data is somehow abridged. Thus, user can interact in utmost haste.
Step 7: When data is successfully loaded, the empty folder should be returned to the Staging Area as acknowledgement of successful loading.
Step 8: While accessing the Marts, it must be specified which Mart is requisite for the sake of reference and retrieval.

6 Conclusion

The section focuses on the fact that the amalgamation of security modus operandi amid data warehouse architecture accentuates the significance of the same and ascertains the righteousness towards the measures adopted for its unsullied implementation and incorporation, ensuring utmost ease in data access and retrieval.

The design and implementation of Double-Ended Bottom-Up Approach of Data Warehouse Architecture Employing Country ISO Code Engendered Cryptographic Algorithm for secured transmission and performance enhancement have been instituted with exclusive prominence for improved intrusion prevention.

References

1. Pal, B., Chowdhury, R., Chatterjee, P., Dasgupta, S., De, M.: Performance enhancement of data warehouse using minimization of query processing proposal to improve ROI. Int. J. Appl. Innov. Eng. Manag. **4**(12), 35–42 (2015). ISSN: 2319 4847, Indexed in Thomson Reuters–Science Citation Index, Indexed in Scopus, Indexed in Thomson Reuters Researcher ID–N–6095–2015
2. Chowdhury, R., Datta, S., Dasgupta, S., De, M.: Implementation of central dogma based cryptographic algorithm in data warehouse for performance enhancement. Int. J. Adv. Comput. Sci. Appl. **6**(11), 29–34 (2015). ISSN (Online): 2156 5570, ISSN (Print): 2158 107X, Included and Indexed in Thomson Reuters Master Journal List with Emerging Sources Citation Index
3. Chowdhury, R., Dey, K.S., Datta, S., Shaw, S.: Design and implementation of proposed drawer model based data warehouse architecture incorporating DNA translation cryptographic algorithm for security enhancement. In: Proceedings of International Conference on Contemporary Computing and Informatics, IC3I 2014, Organized by Sri Jayachamarajendra College of Engineering, Mysore, Proceedings in USB: CFP14AWQ-USB, pp. 55–60. IEEE Digital Xplore (2014). ISBN (10): 1-4799-6628-8, ISBN (13): 978-1-4799-6629-5
4. Chowdhury, R., Chatterjee, P., Mitra, P., Roy, O.: Design and implementation of security mechanism for data warehouse performance enhancement using two tier user authentication techniques. Int. J. Innov. Res. Sci. Eng. Technol. **3**(6), 165–172 (2014). ISSN (Online): 2319 8753, ISSN (Print): 2347 6710
5. Chowdhury, R., Pal, B., Ghosh, A., De, M.: A data warehouse architectural design using proposed pseudo mesh schema. In: Proceedings of First International Conference on Intelligent Infrastructure, CSI ICII 2012, 47th Annual National Convention of Computer Society of India, Science City Auditorium, Kolkata, Tata McGraw Hill Education Private Limited, pp. 138–141 (2012). ISBN (13): 978-1-2590-6170-7, ISBN (10): 1-2590-6170-1
6. Chowdhury, R., Pal, B.: Proposed hybrid data warehouse architecture based on data model. Int. J. Comput. Sci. Commun. **1**(2), 211–213 (2010). ISSN: 0973 7391
7. Saurabh, A.K., Nagpal, B.: A survey on current security strategies in data warehouses. Int. J. Eng. Sci. Technol. **3**(4), 3484–3488 (2011). ISSN: 0975 5462
8. Santos, R.J., Bernardino, J., Vieira, M.: A survey on data security in data warehousing: issues, challenges and opportunities. In: International Conference on Computer as a Tool, EUROCON IEEE, Lisbon, pp. 1–4. IEEE Digital Xplore (2011). ISBN: 978-1-4244-7486-8
9. Vieira, M., Vieira, J., Madeira, H.: Towards data security in affordable data warehouse. In: 7th European Dependable Computing Conference (2008)
10. Gosain, A., Arora, A.: Security issues in data warehouse: a systematic review. Proc. Comput. Sci. **48**, 149–157 (2015). ISSN: 1877 0509. Proceedings of International Conference on Intelligent Computing, Communication and Convergence, ICCC 2015, Organized by Interscience Institute of Management and Technology, Bhubaneswar, Odisha
11. Chaudhuri, S., Dayal, U.: An overview of data warehousing and OLAP technology. Newsl. ACM SIGMOD Rec. **26**(1), 65–74 (1997)
12. Farhan, M.S., Marie, M.E., El-Fangary, L.M., Helmy, Y.K.: An integrated conceptual model for temporal data warehouse security. Comput. Inf. Sci. **4**(4), 46–57 (2011)
13. Golfarelli, M., Rizzi, S.: A methodological framework for data warehouse design. In: Proceedings of ACM First International Workshop on Data Warehousing and OLAP, DOLAP, Washington, pp. 3–9 (1998)
14. Chowdhury, R., Bose, R., Sengupta, N., De, M.: Logarithmic formula generated seed based cryptographic technique using proposed alphanumeric number system and Rubik rotation algorithm. In: Proceedings of IEEE 2012 International Conference on Communications, Devices and Intelligent Systems, CODIS 2012, Organized by Jadavpur University, Kolkata, Proceedings in CD: IEEE Catalog Number–CFP1207U-CDR. ISBN: 978-1-4673-4698-6,

Proceedings in Print: IEEE Catalog Number–CFP1207U-PRT, ISBN: 978-1-4673-4697-9, pp. 564–567 (2012). IEEE Digital Xplore, ISBN: 978-1-4673-4700-6

15. Chowdhury, R., Ghosh, S., De, M.: String graphixification based asymmetric key cryptographic algorithm using proposed concepts of GDC and S-loop matrix. In: Proceedings of IEEE/OSA/IAPR International Conference on Informatics, Electronics & Vision 2012, ICIEV 2012, Organized by University of Dhaka, Dhaka, Bangladesh, Proceedings in CD: IEEE Catalog Number: CFP1244S-CDR, ISBN: 978-1-4673-1152-6, Proceedings in Print: IEEE Catalog Number–CFP1244S-PRT, ISBN: 978-1-4673-1151-9, Conference Proceedings: ISSN: 2226 2105, pp. 1152–1157. IEEE Digital Xplore. ISBN: 978-1-4673-1153-3

Factors Affecting Crime Against Women Using Regression and K-Means Clustering Techniques

Bhajneet Kaur, Laxmi Ahuja and Vinay Kumar

Abstract The basic meaning of crime against women is direct or indirect mental or physical torture or cruelty towards women. Crime against women is increasing every year and as per the research they have doubled over the past ten years, according to latest data released by the NCRB (National Crime Records Bureau). As many as 2.24 million approx. crimes were reported against women over the past decade. On an average 25 crime per hour against women are reported, at least a complaint every two minutes. To control crime, the eyes have to be set on the factors which are influencing the crime against women. For this consideration there are various factors affecting the crime against women. In this paper factors are identified for crime against women. The impact of the individual factor has been checked for the overall crime rate in Delhi on the basis of regression analysis using SPSS tool and thereafter K-means clustering technique has been applied to classify the respondents or cases into clusters on the basis of degree of crime rate for various factors influencing the crime against women.

Keywords Crime against women · Regression analysis · K-means clustering · SPSS

B. Kaur (✉) · L. Ahuja
AIIT, Amity University, Noida, UP, India
e-mail: bhajneetahuja@gmail.com

L. Ahuja
e-mail: lahuja@amity.edu

V. Kumar
Vivekananda Institute of Professional Studies, GGSIPU, New Delhi, India
e-mail: vinay5861@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_15

1 Introduction

By definition crime against women, refers to the crime committed against girls and women. It is a crime which is increasing rapidly now a day in various parts of the country. The main purpose of this research is to analyze the factors that are affecting crime against women and to make positive prevention. In recent years, the absolute number and relative ratio of crime against women has increased per year which cannot be ignored. The crime against women has seriously affected the married lives, dignity and social status of a woman. In this context, it is particularly necessary to understand and analyze the factors affecting crime and bring forward the corresponding measures to control it [1].

According to National Crime Record Bureau Report, under the sections of SLL and IPC there were 337922 cases reported in India during 2014 for crime against women. But in 2013 there were total 309546 cases. Thus during 2014 it was found that 9.2% crime rate has been increased. Even these types of crimes have increased rapidly during 4 years spam 2010 to 2014 with 213585 cases were reported in 2010, which had increased to 228649 cases in 2011, which further had increased to 244270 number of cases in the year of 2012 and 309546 number of cases in the year of 2013. In 2014, there were total 337922 number of cases were reported. The crime rate was calculated corresponding number of crimes against women as 56.3 in the year of 2014. In Delhi, 169.1 crime rates has been calculated, which is the highest rate as compared to 56.3 of overall country rate (India) level during 2014: [2].

2 Related Study

A number of studies have been conducted for the factors and causes which are influencing crimes against women. SK Khandelwal explained the causes and consumption of women crime in the 2nd chapter of his PHD thesis title "A Socio legal Study of Crime Against Women -Critical Reviews of Protective Law", mainly focused on social factors, personal and psychological factors like poverty and some legal causes: [3]. There are various risk factors of crimes against women are concluded in the report *Ending Violence against Women and Girls Programming Essentials*. As mentioned in the report there are variety of factors based upon relationship, individual, society and community intersect to increase the risk and rate of crimes against women and girls. Conflict and tension within an intimate partner relationship or marriage is one of the main factors for crime against women. Other factors are male control over decision making and assets (Male Ego), limited economic opportunities, lack of safe spaces for women and girls, low levels of awareness of law enforcement, lack of punishment etc.: [4]. Though the study is not able to address the criteria of causal inferences, it is consistent with the hypothesis that it is the intimate partner violence that contributes to mental health problems in women. It is these issues that need careful attention: [5].

3 Objective Study

(a) To identify the factors affecting the Crime against women.
(b) To validate the factors for finding the impact of each one on overall crime rate against women by applying regression technique.
(c) To generate the ranking of factors.
(d) To classify the respondents into 3 clusters on the basis of degree of crime rate for various factors influencing the crime against women by using K-means clustering technique.
(e) To discuss implication of this research and suggest directions for future research.

4 Identification and Significance of Factors Affecting the Crime Against Women

On the basis of the literature review, discussion with experts and through the brainstorming session, 11 main factors are identified (i) Male Ego (ii) Gender Discrimination [6] (iii) Electronic media and social media [7] (iv) Lack of knowledge about Act and Statute (v) Poverty (vi) Inappropriate presentation of women characters on TV or silver screens (vii) Lack of safe spaces for women and girls (viii) Poor Upbringing (ix) Lack of Punishment (x) Conflict and tension within intimate partner relationship and marriage (xi) Lack of Implementations of government initiatives. The primary dataset was collected using Questionnaire-based survey method. The majority of respondents for the survey consist of women and girls. The purpose of this survey was to identify the impact of various identified factors on crime rate against women in different regions of Delhi. Significance of each factor has been tested using Regression Model.

5 Validation of Factors

5.1 Data Normality

In order to check the normality of data, Shapiro-Wilk test was used. After applying Shapiro-Wilk test the result has been generated in the given table. As shown in the Table 1 the significance value or p value is greater than 0.05 in all the cases, thus it indicates that the data is normal and further statistical tests can be used.

Table 1 Data normality test

Data	Shapiro_Wilk	
	DF	Sig
Male Ego	407	**0.066**
Gender discrimination	407	**0.072**
Electronic media and social media	407	**0.080**
Lack of knowledge about Act and statute	407	**0.091**
Poverty	407	**0.062**
Inappropriate presentation of women characters on TV or silver screens	407	**0.077**
Lack of safe spaces for women and girls	407	**0.081**
Poor upbringing	407	**0.063**
Lack of punishment	407	**0.073**
Conflict and tension within intimate partner relationship and marriage	407	**0.083**
Lack of implementations of government initiatives	407	**0.064**
Crime rate	407	**0.075**

5.2 Hypothesis Formulation of Factors

For each factor Null-Hypothesis is formulated as follows:

H(0a) There is no significant impact of Male Ego on crime rate against women.

H(0b) There is no significant impact of Gender Discrimination on crime rate against women.

H(0c) There is no significant impact of Electronic media and social media on crime rate against women.

H(0d) There is no significant impact of Lack of knowledge about Act and Statute on crime rate against women.

H(0e) There is no significant impact of Poverty on crime rate against women.

H(0f) There is no significant impact of inappropriate presentation of women characters on TV or silver screens on crime rate against women.

H(0g) There is no significant impact of Lack of safe spaces for women and girls on crime rate against women.

H(0h) There is no significant impact of Poor Upbringing on crime rate against women.

H(0i) There is no significant impact of Lack of Punishment on crime rate against women.

H(0j) There is no significant impact of Conflict and tension within intimate partner relationship and marriage on crime rate against women.

H(0k) There is no significant impact of Lack of Implementations of government initiatives on crime rate against women.

5.3 Regression Model for Each Factor

Hypothesis for each factor is validated using Regression model. The purpose of regression model is to find the impact of each independent factor over the dependent factor i.e. crime rate. Significance of each factor has been checked on the basis of p value and percentage of impact by each factor has been measured through Adjusted R Square.

Impact of Male Ego on Crime rate. Null Hypothesis: H(0a) There is no significant impact of Male Ego on crime rate against women (Table 2).

Since the Adjusted R square is found to be 0.100 which indicates that 10% impact of Male Ego is explained the overall crime rate against women in Delhi. The significant value is found to be 0.000 (highly significant) which is less than 0.05. Thus, null hypothesis (H0) is rejected and alternate hypothesis (H1) is accepted. So, there is a significant impact of Male Ego on overall crime rate against women.

Impact of Gender Discrimination on Crime rate. Null Hypothesis: H(0b) There is no significant impact of Gender Discrimination on crime rate against women (Table 3).

Since the Adjusted R square is found to be 0.018 which indicates that 1.8% impact of Gender Discrimination explained the overall crime rate against women in Delhi. The significant value is found to be 0.004 which is less than 0.05. Thus, null hypothesis (H0) is rejected and alternative hypothesis (H1) is accepted. So, there is a significant impact of Gender Discrimination on overall crime rate against women.

Impact of Electronic media and social media on Crime rate. Null Hypothesis: H0c) There is no significant impact of Electronic media and social media on crime rate against women (Table 4).

Since the Adjusted R square is found to be 0.078 which indicates that 7.8% impact of Electronic media and social media explained the overall crime rate against women in Delhi. The significant value is found to be 0.000 (highly significant) which is less than 0.05. Thus, null hypothesis (H0) is rejected and alternative hypothesis (H1) is accepted. So, there is a significant impact of Electronic media and social media on overall crime rate against women.

Table 2 Regression analysis for the factor Male Ego on crime rate

R	R_Square	Adjusted R^2	Std_Error_of_Estimate	Change statistic				
				R_square_Change	F_Change	Df_1	Df_2	Sig_F change
0.320	0.103	0.100	0.657	0.103	46.337	1	405	0.000

Table 3 Regression analysis for the factor gender discrimination on crime rate

R	R_Square	Adjusted R^2	Std_Error_of_Estimate	Change statistic				
				R_square_Change	F_Change	Df_1	Df_2	Sig_F change
0.143	0.020	0.018	0.687	0.020	8.446	1	405	0.004

Table 4 Regression analysis for the factor electronic media and social media on crime rate

R	R_Square	Adjusted R^2	Std_Error_ of_Estimate	Change statistic				
				R_square_ Change	F_Change	Df_1	Df_2	Sig_F change
0.284	0.080	0.078	0.665	0.080	35.436	1	405	0.000

Table 5 Regression analysis for the factor lack of knowledge about act and statute on crime rate

R	R_Square	Adjusted R^2	Std_Error_ of_Estimate	Change statistic				
				R_square_ Change	F_Change	Df_1	Df_2	Sig_F change
0.347	0.120	0.118	0.651	0.120	55.368	1	405	0.000

Table 6 Regression analysis for the factor poverty on crime rate

R	R_Square	Adjusted R^2	Std_Error_ of_Estimate	Change statistic				
				R_square_ Change	F_Change	Df_1	Df_2	Sig_F change
0.206	0.042	0.040	0.679	0.042	17.894	1	405	0.000

Impact of Lack of knowledge about Act and Statute on Crime rate. Null Hypothesis: H(0d) There is no significant impact of Lack of knowledge about Act and Statute on crime rate against women (Table 5).

Since the Adjusted R square is found to be 0.118 which indicates that 11.8% impact of Lack of knowledge about Act and Statute explained the overall crime rate against women in Delhi. The significant value is found to be 0.000 (highly significant) which is less than 0.05. Thus, null hypothesis (H0) is rejected and alternative hypothesis (H1) is accepted. So, there is a significant impact of Lack of knowledge about Act and Statute on overall crime rate against women.

Impact of poverty on Crime rate. Null Hypothesis: H(0e) There is no significant impact of Poverty on crime rate against women (Table 6).

Since the Adjusted R square is found to be 0.040 which indicates that 4% impact of Poverty explained the overall crime rate against women in Delhi. The significant value is found to be 0.000 (highly significant) which is less than 0.05. Thus, null hypothesis (H0) is rejected and alternative hypothesis (H1) is accepted. So, there is a significant impact of Poverty on overall crime rate against women.

Impact of Inappropriate presentation of women characters on TV or silver screens on Crime rate. Null Hypothesis: H(0f) There is no significant impact of Inappropriate presentation of women characters on TV or silver screens on crime rate against women (Table 7).

Since the Adjusted R square is found to be 0.057 which indicates that 5.7% impact of inappropriate presentation of women characters on TV or silver screens

Table 7 Regression analysis for the factor inappropriate presentation of women characters on TV or silver screens on crime rate

R	R_Square	Adjusted R^2	Std_Error_ of_Estimate	Change statistic				
				R_square_ Change	F_Change	Df_1	Df_2	Sig_F change
0.244	0.059	0.057	0.673	0.059	25.591	1	405	0.000

Table 8 Regression analysis for the factor inappropriate presentation of women characters on TV or silver screens on crime rate

R	R_Square	Adjusted R^2	Std_Error_ of_Estimate	Change statistic				
				R_square_ Change	F_Change	Df_1	Df_2	Sig_F change
0.995	0.990	0.990	0.070	0.990	39189.483	1	405	0.000

explained the overall crime rate against women in Delhi. The significant value is found to be 0.000 (highly significant) which is less than 0.05. Thus, null hypothesis is rejected and alternative hypothesis is accepted. So, there is a significant impact of inappropriate presentation of women characters on TV or silver screens on overall crime rate against women.

Impact of Lack of safe spaces for women and girls on Crime rate. Null Hypothesis: H(0g) There is no significant impact of Lack of safe spaces for women and girls on crime rate against women (Table 8).

Since the Adjusted R square is found to be 0.040 which indicates that 99% impact of Lack of safe spaces for women and girls explained the overall crime rate against women in Delhi. The significant value is found to be 0.000 (highly significant) which is less than 0.05. Thus, null hypothesis (H0) is rejected and alternative hypothesis (H1) is accepted. So, there is a significant impact of Lack of knowledge about Act and Statute on overall crime rate against women.

Impact of Poor Upbringing on Crime rate. Null Hypothesis: H(0h) There is no significant impact of Poor Upbringing on crime rate against women (Table 9).

Since the Adjusted R square is found to be 0.955 which indicates that 95.5% impact of Poor Upbringing explained the overall crime rate against women in Delhi. The significant value is found to be 0.000 (highly significant) which is less than 0.05. Thus, null hypothesis is rejected and alternative hypothesis is accepted. So, there is a significant impact of Poor Upbringing on overall crime rate against women.

Table 9 Regression analysis for the factor poor upbringing on crime rate

R	R_Square	Adjusted R^2	Std_Error_ of_Estimate	Change statistic				
				R_square_ Change	F_Change	Df_1	Df_2	Sig_F change
0.977	0.955	0.955	0.147	0.955	8617.740	1	405	0.000

Table 10 Regression analysis for the factor lack of punishment on crime rate

R	R_Square	Adjusted R^2	Std_Error_ of_Estimate	Change statistic				
				R_square_ Change	F_Change	Df_1	Df_2	Sig_F change
0.118	0.014	0.011	0.689	0.014	5.674	1	405	0.018

Impact of Lack of Punishment on Crime rate. Null Hypothesis: H(0i) There is no significant impact of Lack of Punishment on crime rate against women (Table 10).

Since the Adjusted R square is found to be 0.011 which indicates that 1.1% impact of Lack of Punishment explained the overall crime rate against women in Delhi. The significant value is found to be 0.018 (significant) which is less than 0.05. Thus, null hypothesis is rejected and alternative hypothesis is accepted. So, there is a significant impact of Lack of Punishment on overall crime rate against women.

Impact of Conflict and tension within intimate partner relationship and marriage on Crime rate. Null Hypothesis: H(0j) There is no significant impact of Conflict and tension within intimate partner relationship and marriage on crime rate against women (Table 11).

Since the Adjusted R square is found to be 0.916 which indicates that 91.6% impact of Conflict and tension within intimate partner relationship and marriage explained the overall crime rate against women in Delhi. The significant value is found to be 0.000 (highly significant) which is less than 0.05. Thus, null hypothesis is rejected and alternative hypothesis is accepted. So, there is a significant impact of Conflict and tension within intimate partner relationship and marriage on overall crime rate against women.

Impact of Lack of Implementations of government initiatives on Crime rate. Null Hypothesis: H(0k) There is no significant impact of Lack of Implementations of government initiatives on crime rate against women (Table 12).

Table 11 Regression analysis for the factor conflict and tension within intimate partner relationship and marriage on crime rate

R	R_Square	Adjusted R^2	Std_Error_ of_Estimate	Change statistic				
				R_square_ Change	F_Change	Df_1	Df_2	Sig_F change
0.957	0.916	0.916	0.201	0.916	4444.047	1	405	0.000

Table 12 Regression analysis for the factor lack of Implementations of government initiatives on crime rate

R	R_Square	Adjusted R^2	Std_Error_ of_Estimate	Change statistic				
				R_square_ Change	F_Change	Df_1	Df_2	Sig_F change
0.342	0.117	0.114	0.652	0.117	53.474	1	405	0.000

Table 13 Ranking of factors

Descriptive_Statistics of factors						
	N (SS)	Mini	Max	Mean	Std deviation	Ranking of factors
Male Ego	407	2	5	3.57	0.824	**11**
Gender discrimination	407	3	5	4.14	0.414	**6**
Electronic media and social media	407	3	5	4.19	0.562	**5**
Lack of knowledge about Act and statute	407	2	5	3.97	0.823	**7**
Poverty	407	1	5	3.85	0.970	**9**
Inappropriate presentation of women characters on TV or silver screens	407	2	5	4.22	0.741	**4**
Lack of safe spaces for women and girls	407	2	5	4.30	0.693	**1**
Poor upbringing	407	2	5	4.28	0.695	**3**
Lack of punishment	407	1	5	3.58	0.981	**10**
Conflict and tension within intimate partner relationship and marriage	407	2	5	4.29	0.704	**2**
Lack of implementations of government initiatives	407	2	5	3.95	0.869	**8**

Since the Adjusted R square is found to be 114 which indicates that 11.4% impact of Lack of Implementations of government initiatives explained the overall crime rate against women in Delhi. The significant value is found to be 0.000 (highly significant) which is less than 0.05. Thus, null hypothesis (H0) is rejected and alternate hypothesis (H1) is accepted. So, there is a significant impact of Lack of Implementations of government initiatives on overall crime rate against women.

6 Ranking of Factors

As shown in the Table 13, Ranking of factors were obtained based on their Mean value. The Mean value of the factor **Lack of safe spaces for women and girls** is highest among all other values. So, for this reason it has been given highest rank (Rank-1). On the basis of same procedure other rankings have been obtained.

7 K-Means Clustering Technique

The method of clustering observations into specific number of disjoint clusters or homogeneous groups is called K-means. K is considered as the number of specified clusters. To determine which observation is needed to append to which cluster is measured by the various distances calculations. The aim of this algorithm is to minimize the measures between the given observation and center of the cluster by

Table 14 Initial output of cluster centers

ZScore: factors	Cluster		
	1	2	3
Zscore: Male Ego	−0.68851	−1.90160	0.52458
Zscore: Gender discrimination	−0.32646	−0.32646	2.08933
Zscore: Electronic media and social media	1.44160	−2.11435	1.44160
Zscore: Lack of knowledge about act and statute	−1.18286	−2.39857	1.24857
Zscore: poverty	−1.90771	1.18567	−2.93884
Zscore: Inappropriate presentation of women characters on TV or silver screens	−0.29176	−2.99054	1.05763
Zscore: Lack of safe spaces for women and girls	−3.31461	−0.42895	1.01388
Zscore: Poor upbringing	−3.28105	−0.40306	1.03593
Zscore: Lack of punishment	1.44464	−0.59338	0.42563
Zscore: Conflict and tension within intimate partner relationship and marriage	−3.24803	−0.40513	1.01632
Zscore: Lack of implementations of government initiatives	−2.23851	0.06218	0.06218

number of iterations appending an observation to any cluster and iterations are terminated on achieving the lowest distance measure [8].

Common distance measures include the Euclidean distance, the Euclidean squared distance and the Manhattan or City distance.

The Euclidean measure corresponds to the shortest geometric distance between two points: [9].

$$d = \sqrt{\sum_{i=1}^{N} (x_i - y_i)} \tag{1}$$

A faster way of determining the distance is by use of the squared Euclidean distance which calculates the above distance squared, i.e.

$$d_{zq} = \sqrt{\sum_{i=2}^{N} (x_i - y_i)^z} \tag{2}$$

In the below Table 14, Initial cluster Centers have been calculated for every factor for the segregation of respondents for crime rate. The value of k is taken as 3 so the overall response of crime rate is divided into 3 clusters using k-means clustering algorithm by using SPSS tool. The whole process of K-means clustering is done through Z-score descriptive of each factor. Z-score is used in the k-means clustering to standardize the values of features from different range (dynamic) into specific range. Z-score standardization is more effective and efficient than other standardization methods: [10, 11] (Table 15).

Table 15 Iteration history

Iterations	Changes in cluster centers		
	1	2	3
1	3.122	3.878	3.924
2	0.000	0.419	0.246
3	0.000	0.055	0.038
4	0.000	0.000	0.000

Table 16 Final output of cluster centers

ZScore: factors	Cluster		
	1	2	3
Zscore: Male Ego	−0.68851	−0.66473	0.56300
Zscore: Gender discrimination	−0.32646	−0.50014	0.39500
Zscore: Electronic media and social media	0.84894	−0.44096	0.17852
Zscore: Lack of knowledge about act and statute	−0.77762	−0.49157	0.45643
Zscore: poverty	−0.53288	−0.37787	0.34117
Zscore: Inappropriate presentation of women characters on TV or silver screens	0.60783	−0.85621	0.50200
Zscore: Lack of safe spaces for women and girls	−2.35273	−0.34408	0.58952
Zscore: Poor upbringing	−2.32172	−0.32782	0.57363
Zscore: Lack of punishment	1.10497	−0.50680	0.18586
Zscore: Conflict and tension within intimate partner relationship and marriage	−2.30040	−0.30293	0.55322
Zscore: Lack of implementations of government initiatives	−0.32127	−0.51675	0.40572

Firstly the initial cluster centers have been measured for every factor on the basis of Z-score descriptive as shown in the Table for all the responses.

Then four iterations have been made to change the cluster center because in the fourth iteration all the iterations are found as same result. On the basis of the iterations, the result of final clusters centers have been generated which is shown in the Table 16 and graphically represented in the Fig. 1.

7.1 Finding and Interpretation

From Fig. 1, it has been found that the overall responses have been divided into three clusters of homogeneous responses based on the degree of impact of each factor. From Table 17, majority of respondents for homogeneous result has been

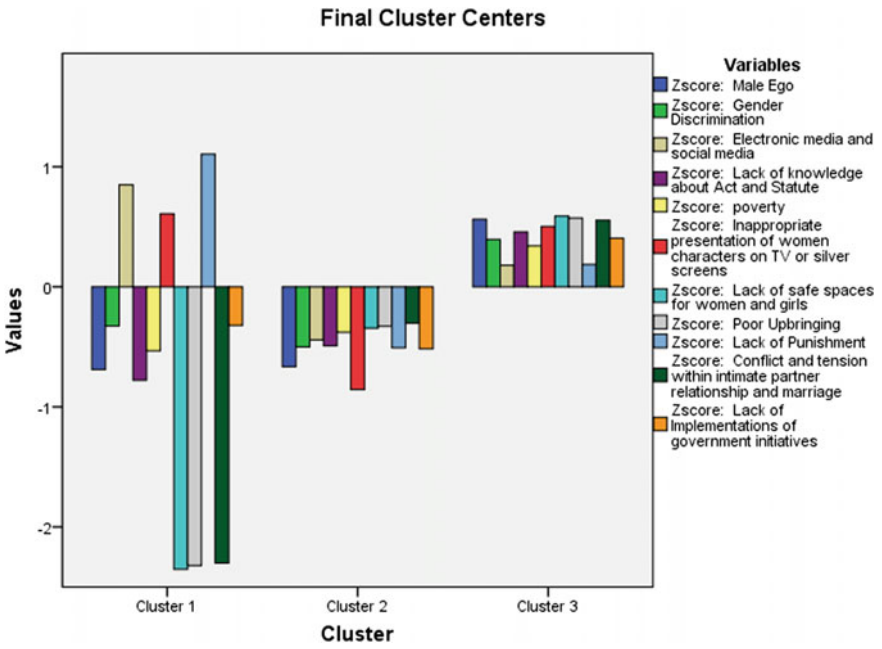


Fig. 1 Final cluster center

Table 17 Number of responses (cases) in each cluster

Cluster	1	33.000
	2	153.000
	3	**221.000**
Valid		407.000
Missing		19.000

found in the cluster-3. This shows that out of 407, 221 respondents have positive result for every factor which has influenced the crime against women. According to the cluster-3, highest value achieved for the factor "Lack of safe spaces for women and girls" it means that this particular factor shows highly impact on the crime rate which is same as formulated in the regression model.

Here is one more information has been displayed in the Table 18 which shows the significance value (P value) for every factor of crime against women is less than 0.05, so the responses by all the respondents have significant by k-means clustering technique.

Table 18 Representation of ANOVA

	Cluster		Error term		F	Sig.
	Mean_Square	Df	Mean_Square	Df		
Zscore: Male Ego	76.650	2	0.625	404	122.542	0.000
Zscore: Gender discrimination	38.135	2	0.816	404	46.725	0.000
Zscore: Electronic media and social media	30.288	2	0.855	404	35.424	0.000
Zscore: Lack of knowledge about act and statute	51.483	2	0.750	404	68.637	0.000
Zscore: poverty	28.470	2	0.864	404	32.952	0.000
Zscore: Inappropriate presentation of women characters on TV or silver screens	90.024	2	0.559	404	160.962	0.000
Zscore: Lack of safe spaces for women and girls	138.792	2	0.318	404	436.642	0.000
Zscore: Poor ppbringing	133.523	2	0.344	404	388.207	0.000
Zscore: Lack of punishment	43.612	2	0.789	404	55.271	0.000
Zscore: Conflict and tension within intimate partner relationship and marriage	128.154	2	0.371	404	345.875	0.000
Zscore: Lack of Implementations of government initiatives	40.320	2	0.805	404	50.066	0.000

8 Conclusion and Future Scope

In this paper, analysis of the factors has been done which are affecting the crime rate against women in different regions of Delhi. The factors have been identified through literature review, Expert opinion and brainstorming sessions. Then a survey has been conducted on these factors for the validation purpose and the primary data has been collected with the help of questionnaire and majority of respondents are women and girls of different age groups. A linear Regression technique has been applied for the analysis of factors and to test the impact of every individual factor on crime rate. As a result of regression model, all the identified factors that have influenced the crime rate with different percentages have been proved hypothetically. Then the ranking has been generated for all the factors on the basis of degree of impact on the crime rate in the Table 13. The highly impacted factor found on crime rate is "Lack of safe spaces for women and girls". Then K-means clustering algorithm has been used to obtain the clusters of homogeneous responses. Three clusters are shown in Fig. 1, it has been observed that results of cluster-3 have majority of responses (221 out of 407) which includes positive result for all the factors. So it concludes that every identified factor has to be considered for the crime rate against women but the most considerable factor is "Lack of safe spaces for women and girls" because this factor

is at the highest point in the cluster-3. The factor also came on the highest ranking also by the regression technique. In future any predictive model can be developed for crime against women on the basis of the ranking of validated factors so that measures can be taken to control the increasing crime against women.

References

1. Parihar, A.: Crime Against Women in Haryana: an Analysis, vol. 4, Issue 11, pp. 16–24, Nov 2015, http://www.ijhssi.org
2. National crime record bureau report: Chapter-5 Crime against Women, http://ncrb.nic.in/StatPublications/CII/CII2014/Compendium%202014.pdf (2014)
3. Khandelwal, S.K.: Crime Against Women: Causes And Compulsions, A Socio-legal Study of Crimes Against Women: A Critical Review of Protective Law, chapter-2, http://shodhganga.inflibnet.ac.in/bitstream/10603/33145/2/chapter%202.pdf (2015)
4. UN Women: Ending violence against Women and Girls: Progamming essentials, June 2013, http://www.endvawnow.org/uploads/modules/pdf/1372349234.pdf (2013)
5. Golding, J.M.: Intimate partner violence as a risk factor for mental disorders: a meta-analysis. J. Fam. violence **14**(2), 99–132, (1999)
6. Morrison, A., et al.: Addressing Gender Based Violence- A Critical Review of Interventions, vol. 22, pp. 25–51. World Bank Research Observer, Washington, DC
7. Malamuth, N.M., Check, J.V.P.: The effects of mass media exposure on acceptance of violence against women: a field experiment. J. Res. Pers. **15**(4), 436–446 (1981)
8. Pham, D.T., Stefan, S.D., Nguyen, C.D.: Selection of K in K-means clustering. Proc Inst. Mech. Eng. Part C: J. Mech. Eng. Sci. **219**(1), 103–119 (2005)
9. The K-means Clustering Algorithm: Chapter-1, pp. 1–6, http://kom.aau.dk/group/04gr742/pdf/kmeans_worksheet.pdf
10. Mohamad, I.B., Usman, D.: Standardization and its effects on K -Means clustering algorithm. Res. J. Appl. Sci. Eng. and Technol. **6**(17), 3299–3303 (2013), http://maxwellsci.com/print/rjaset/v6-3299-3303.pdf
11. Chen, Z.: Exploring new trends of university libraries by SPSS cluster analysis method: Take Wuhan University of Technology as an example. In: 2011 6th International Conference on Product Innovation Management (ICPIM). IEEE (2011)

Energy Efficient Data Gathering in Wireless Sensor Networks Using Rough Fuzzy C-Means and ACO

Sanjoy Mondal, Saurav Ghosh and Pratik Dutta

Abstract Data gathering from inhospitable terrains such as volcanic area, dense forest, sea bed are a major application area of wireless sensor network (WSN). The replacements of sensor node batteries are not feasible and as a result all the protocols in WSN should be energy efficient to elongate network lifetime. In hierarchical routing protocol (HRP) nodes are assigned different tasks of varying energy intensity as per their role which are interchanged across rounds. It leads to load balancing and energy preservation. We propose in this paper an energy efficient load balanced data gathering method based on rough fuzzy c-means (RFCM) and ant colony optimization (ACO) and coin it as RFCM-ACO. The deployed are partitioned into clusters by RFCM followed by ACO-based lower and upper chain formation. The chain leader (CL) for lower chain and super leader (SL) for upper chain are elected using a fuzzy inference system (FIS). Simulation results indicate that RFCM-ACO outperforms LEACH, PEGASIS and Hybrid_FCM in terms of network lifetime and load balance.

Keywords Clustering · Energy efficiency · Load balance · RFCM · Network lifetime · ACO

S. Mondal · S. Ghosh (✉)
A.K. Choudhury School of I.T., University of Calcutta, Kolkata, India
e-mail: sauravghoshcu@gmail.com

S. Mondal
e-mail: sanjoymondal1988@gmail.com

P. Dutta
Department of Computer Science and Engineering, University of Calcutta, Kolkata, India
e-mail: pratikcsc@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_16

1 Introduction

The advancement of transducer and wireless technology has made it feasible for bulk manufacture of robust sensor nodes. They maintain multi-hop connectivity amongst themselves and to the base station (BS) in a wireless manner forming a wireless sensor network (WSN) [1, 2]. A variety of natural parameters like humidity, temperature, pressure are sensed by the sensors and the data is routed to a far away BS where it is analyzed to take appropriate decisions. In a majority of applications, sensor nodes are deployed in hazardous environment where it is not possible to replace node battery. As a consequence, all the protocols designed for WSN must be energy efficient to elongate the network lifetime.

In hierarchical routing protocols (HRP) [3, 4] for WSN's different roles for the nodes are determined depending on their energy requirements. These various roles are assigned to different nodes in a round being interchanged in future rounds for load balancing with an overall objective of reducing energy dissipated per round and subsequent extension of WSN lifetime. HRP's are broadly classified into cluster-based protocols like LEACH [3] and chain based ones like PEGASIS [4]. In cluster-based HRP's nodes are partitioned into clusters with the election of a cluster head (CH). Data flows from the nodes to CH's being aggregated by it and transmitted to the BS. The role of CH is interchanged every round for load balancing. On the other hand in chain-based HRP's, a single chain is constructed out of the deployed nodes with the selection of a chain leader (CL) to transmit the fused data directly to the BS. At the onset of a data transfer round CLs are selected from random positions along the chain for load balancing. In both approaches, there is no strict control over the compactness of the cluster or length of the chain resulting in poor performance.

In this paper, we propose a hybrid HRP coined as RFCM-ACO by combining the characteristics of both approaches. In RFCM-ACO, the deployed nodes are divided into fixed number of clusters using rough fuzzy c-means (RFCM) [5] to get a better degree of control over cluster compactness. RFCM incorporates the concept of rough set to the basic fuzzy c-means (FCM) [6] proposed by Bezdek. The cluster chains referred as lower chains are constructed using ant colony optimization (ACO) [7] to reduce chain length. The selection of chain leader (CL) is based on a fuzzy inference system (FIS) [8] for better load balancing. An upper chain consisting of the CL's is constructed using ACO with the election of a super leader (SL) by the FIS. Simulation results establish our claim that RFCM-ACO performs better over LEACH [3], PEGASIS [4] and Hybrid-FCM [9] in respect to WSN lifetime and load distribution.

The rest of the paper is organized into Sect. 2 giving a survey of related works with our detailed proposal in Sect. 3. Our simulation result appears in Sect. 4 followed by the concluding remarks in Sect. 5.

2 Related Works

A number of HRP's for data gathering in WSN exists in literature of both categories namely cluster based and chain based. LEACH [3] is a pioneering cluster-based HRP proposed by Wendi et al. partitions the deployed nodes into clusters with the election of a cluster head (CH) in a distributed manner. It extends WSN life by a factor of six compared to the naïve direct transmission method. The drawbacks of distributed cluster formation in LEACH are removed in LEACH-C [3] by a centralized one based on simulated annealing. LEACH-C improves upon LEACH by about 20%. On the other hand, PEGASIS [4] is one of the earliest chain-based HRP's in WSN. All the nodes spread across the deployment area are connected together in a single chain in a greedy manner. In different rounds, chain leaders (CL) are selected at random locations along the chain to send the aggregated data to the BS.

A fuzzy c-means (FCM) [6]-based data routing algorithm was proposed by Hadjila et al. [9] (Hybrid-FCM) where nodes are partitioned using FCM with ant colony optimization (ACO) [7]-based cluster chain formation. The CL selection is based on maximum residual energy. Instead of fixing the number of clusters a method determining the optimal cluster number and then using FCM to partition was proposed by Lam and Hrong [10]. The cluster head selection is based on fuzzy logic system with number of nodes in a cluster and their sensing area as inputs.

Chen [11] proposed a cluster-based HRP using FCM algorithm for cluster construction and this infrastructure remains unchanged throughout the network lifetime. The CH is selected based on Q factor which is defined as the ratio of a node residual energy and square of its distance from the BS. In [12] proposed by Chourasia et al., the deployment region is divided into three regions in the form of far, near and close with respect to their distance from the BS. The nodes in the close region send directly to BS while nodes in the other regions send indirectly via appropriate CH's in the intermediate regions. Nayak and Devulapalli [13] proposes an extension to LEACH in the form of introducing the concept of super cluster head (SCH). The basic clustering is similar to LEACH and amongst the CH's a SCH is elected based on fuzzy logic. The CH's sends the aggregated data to the SCH which in turn sends it to the BS. F-MCHEL [14] proposed by T. Sharma et al. is a cluster-based HRP where CHs are elected by the FIS with inputs as residual energy and proximity to the BS. The CH having the maximum residual energy is elected as a Master Cluster Head (MCH) which sends the aggregated data to BS. He and Dai [15] proposed a centralized clustering algorithm GFCM, which combines the concept of Genetic Algorithm (GA) and FCM to produce compact clustering infrastructure remaining fixed throughout the network lifetime. Another centralized partitioning approach is in Decentralized Fuzzy Clustering Protocol (DFCP) [16] where the FCM-based clustering is only done once initially by the BS. In the CH election phase, CH's are elected locally based on a multi-objective function. In TSFL-LEACH [17] proposed a fuzzy logic-based clustering where nodes in a cluster are divided into two types depending on their distance from the CH. A fuzzy

approach is used to elect tentative CH's apart from the main CH. Data is directly send to the main CH by the nodes within its radius while others send to tentative CH's which does a multi-hop relay to the BS. A clustering scheme based on a combination of neural network and fuzzy logic coined as NFEACS was proposed by Julie and Selvi in [18]. The energy level for being a cluster head is estimated by neural network while the fuzzy logic part selects the most appropriate node among a set of candidate nodes. The node mobility is also taken into account. NFEACS performs better when compared with contemporary clustering methods using fuzzy approach. Tomar et al. in [19] proposed a hierarchical routing protocol (HRP) similar to LEACH where cluster head selection is based on fuzzy logic while data routing to the BS from a sensor node is based on ACO. Simulation studies indicate that the proposal performs better than other HRP's in respect to extending WSN lifetime and improving system load balance. A still another extension to LEACH was proposed by Alami and Najid [20] as Energy Efficient Fuzzy Logic Cluster Head (EEFL-CH). The choice of cluster head is based on the FIS with node remaining battery power, proximity to BS and expected efficiency as inputs and results in an extension of system life.

3 Proposed Method

The randomly spread nodes over the deployment area are partitioned by RFCM [5] into fixed number of clusters by the BS. This infrastructure remains fixed throughout the entire lifespan of the WSN eliminating the overhead of cluster formation at the onset of every data transfer round. A PEGASIS like cluster chain is formed using ACO [7] for each cluster and chain leaders (CLs) are elected by the FIS [8] with node's residual energy and distance to the BS as inputs. An upper chain joining the CLs is again constructed by the BS using ACO and a super leader is similarly selected as the CLs. Data flows along the cluster chain being aggregated from node to node till it reaches the CL. Then it moves along the upper chain from CL to CL in a similar way till it reaches the super leader which transmits it directly to the BS. The procedure is outlined by the algorithm RFCM-ACO described below.

Algorithm RFCM-ACO

Step 1. The sensor nodes are randomly deployed across the deployment area.
Step 2. The deployed nodes are partitioned into a fixed number of clusters using RFCM by the BS. The clustering infrastructure remains fixed throughout network lifetime.
Step 3. The nodes in a cluster are linked into a chain using ACO by BS. A chain leader is elected from each such chain based on fuzzy logic with the node's residual energy and its distance from BS as inputs.

Step 4. An upper chain consisting of the chain leader's is constructed using ACO by BS. A super leader is elected based on fuzzy logic with the chain leader's residual energy and its distance from BS as inputs.
Step 5. Data flows along the cluster chain to the chain leader and after that it flows along the upper chain to the super leader. Finally the super leader sends the data to the BS.
Step 6. If the number of alive nodes is equal to zero then stop otherwise go to Step 3.

3.1 Cluster Formation Using RFCM

The concept of crisp clustering maps an object to one and only one cluster. Hard-c-means (HCM) is an example of crisp clustering where object similarity in clusters and non-overlapping nature of clusters are predominant. When an object belongs to number of clusters with varying degree, i.e. overlapping clusters the use of FIS in clustering is more beneficial to conventional one like HCM by allowing the concept of gradual membership. The properties like uncertainty, vagueness and incompleteness are well represented in the rough set theory which can be combined with the FIS to produce compact clusters.

Maji and Pal [5] proposed rough fuzzy c-means (RFCM) which combines the concepts of both fuzzy and rough sets where lower region is crisp in nature while boundary region is fuzzy. RFCM proposed the concept of adding fuzzy membership of fuzzy sets and lower and upper approximation from rough sets. RFCM divides set of n objects into c clusters by minimizing the objective function (1).

$$J_{RFCM} = \begin{cases} w \times \widetilde{A}_1 + w' \times \widetilde{B}_1 & \text{if } \underline{A}(\beta_i) \neq \varphi, B(\beta_i) \neq \varphi \\ \widetilde{A}_1 & \text{if } \underline{A}(\beta_i) \neq \varphi, B(\beta_i) = \varphi \\ \widetilde{B}_1 & \text{if } \underline{A}(\beta_i) = \varphi, B(\beta_i) \neq \varphi \end{cases} \tag{1}$$

$$\widetilde{A}_1 = \sum_{i=1}^{c} \sum_{x_j \in \underline{A}(\beta_i)} (\mu_{ij})^{m_1'} \left\| x_j - v_i \right\|^2; \; \widetilde{B}_1 = \sum_{i=1}^{c} \sum_{x_j \in B(\beta_i)} (\mu_{ij})^{m_1'} \left\| x_j - v_i \right\|^2;$$

$$\mu_{ij} = \left(\sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{kj}} \right)^{\frac{2}{m_1' - 1}} \right)^{-1}; \quad d_{ij}^2 = \left\| x_j - v_i \right\|^2$$

The parameters w and $w'(=1-w)$ correspond to the relative importance of lower and boundary region, respectively, μ_{ij} is the membership value for the jth object in ith cluster, m_1' is the fuzzifier (generally chosen as >1).

In RFCM, each cluster has a centroid, a crisp lower region and a fuzzy boundary region with the lower region affecting the fuzziness of final partitioning. If an object $x_j \in \underline{A}(\beta_i)$ then $x_j \in \underline{A}(\beta_k)$, $\forall k \neq i$ and $x_j \in B(\beta_i), \forall i$. i.e. the object x_j contained in β_i definitely. The weight parameter w is associated with the objects of lower region and in our simulation we choose w to be 0.95 as a weightage to the crispness of objects. On the other hand if $x_j \in B(\beta_i)$ then the object x_j possibly belongs to β_i and several clusters. Hence the objects in boundary region have different and less influence on the centroids and clusters. In RFCM, the lower region is crisp means $\mu_{ij} = 1$, so the value of $\widetilde{A}_1$ in objective function revised as

$$\widetilde{A}_1 = \sum_{i=1}^{c} \sum_{x_j \in \underline{A}(\beta_i)} \left\| x_j - v_i \right\|^2 \tag{2}$$

3.2 Chain Formation Using Ant Colony Optimization

Ant Colony Optimization (ACO) [7] provides a good approximate solution to travelling salesman problem (TSP). The construction of cluster chain and upper chain is similar in nature to TSP and advocates the use of ACO for length optimization. A number of artificial ants are randomly placed at different nodes and all tries to simultaneously find a tour of the cluster nodes or CL's. The *k*th ant placed at node i moves to node j according to the pseudo-random probability defined in Eq. (3),

$$p_{ij}^{k} = \frac{(\tau_{ij})^{\alpha}(\eta_{ij})^{\beta}}{\sum_{l \in N_i^k} (\tau_{il})^{\alpha}(\eta_{il})^{\beta}} \quad \text{if } j \in N_i^k \tag{3}$$

where τ_{ij} and η_{ij} are the pheromone levels and length of the edge (i, j). $\alpha > 0$ *and* $\beta > 0$ are two parameters determining the relative weightage given to pheromone level and distance, respectively. This movement is known as iteration and each ant after a fixed number of iterations completes its tour. At the completion of tour by all the ants the global best tour is found out and the edges belonging to that tour receive pheromone reinforcement as Eq. (4) while pheromone evaporates on others as Eq. (5).

$$\tau_{ij} = \tau_{ij} + (L^{-1}) \tag{4}$$

$$\tau_{ij} = (1 - \rho)\tau_{ij} \tag{5}$$

L is the length of the global best tour and ρ is the pheromone decay factor. The process repeats until between two successive passes there is insignificant change in tour length.

Table 1 Fuzzy rule base table

Rule	Residual energy	Distance	Chance
1	Low	Near	Medium
2	Low	Moderate	Low
3	Low	Far	Low
4	Medium	Near	Little high
5	Medium	Moderate	Medium
6	Medium	Far	Low
7	High	Near	High
8	High	Moderate	Little high
9	High	Far	Medium

3.3 Chain Leader and Super Leader Election Using Fuzzy Logic

The cluster chain leader for each cluster chain and the super leader for the upper chain are selected by the FIS [8] with node residual energy and its distance from the BS as inputs. The FIS comprises of four basic processes namely fuzzification, inference engine, fuzzy rule base and defuzzification. The crisp inputs like residual battery energy of a node and its distance from the BS are changed into fuzzy sets during the fuzzification process. We define three fuzzy levels each for residual node energy and its distance from BS as shown in Table 1. A fuzzy rule base in the form of IF-THEN rules are defined in Table 1 with the output fuzzy variable Chance. The inference engine takes fuzzy input variables and searches the fuzzy rule base table to find a matching value for fuzzy output variable which is then converted into a crisp value by the defuzzification process. The node with the highest chance value is selected as the chain leader or super leader as the case might be.

4 Simulation Environment and Results

A random spread of 100 homogeneous static sensor nodes over a 100 m * 100 m deployment area is considered. The BS is fixed and located at (50,150). We assume that any node can communicate directly with any other node and with the BS in line with [3, 4]. It can vary its transmitting power accordingly. A node dissipates energy in transmission as per the free space or multi-path model depending on the threshold distance d_0 defined as square root of $\varepsilon_{fs}/\varepsilon_{mp}$. The Eqs. (6) and (7) model the loss for transmitting a m bit message over a distance d while Eq. (8) accounts for the loss in the reception of a m bit message as in [3, 4].

Initial node energy of 1 J, packet size of 2000 bits, no obstacles with Omni-directional antenna type are assumed.

$$E_{Tx}(m,d) = E_{elec} * m + \varepsilon_{fs} * m * d^2 \quad \text{if } d < d_0 \tag{6}$$

$$= E_{elec} * m + \varepsilon_{mp} * m * d^4 \quad \text{if } d \geq d_0 \tag{7}$$

$$E_{Rx}(m) = E_{elec} * m \tag{8}$$

In our simulation, we set E_{elec} (in nano Joules/bit) to 50, ε_{fs} (in pico Joules/bit/m^2) to 10, ε_{mp} (in pico Joules/bit/m^2) to 0.0013 and E_{DA} (in nano Joules/bit/signal) to 0.0013. The protocols LEACH, PEGASIS, Hybrid-FCM and RFCM-ACO are simulated using MATLAB.

The number of alive nodes against round is plotted in Fig. 1 for all the protocols for a random deployment of 100 nodes. The corresponding values of First Node Dies (FND), Half of Node Die (HND) and Last Node Dies (LND) in rounds are shown by the bar graph in Fig. 2. It is evident from the results that RFCM-ACO extends WSN lifespan specially the FND which accounts for the full functionality of the network. A good load balance among the nodes is provided as is evident from the fact that nodes die in a narrow time span. The average number of data packets received at BS per round is the throughput indicating the effectiveness and quality of the sensed data. It can be inferred from Fig. 3 that RFCM-ACO gives better throughput over the others.

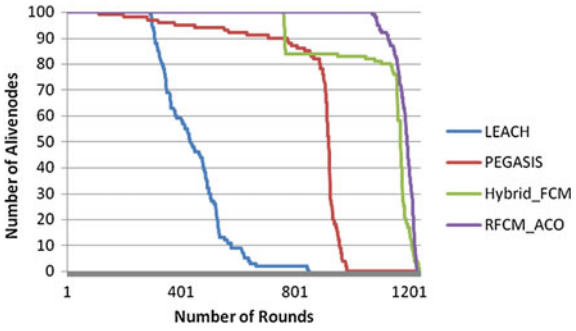


Fig. 1 Number of alive nodes against rounds for a random deployment of 100 nodes

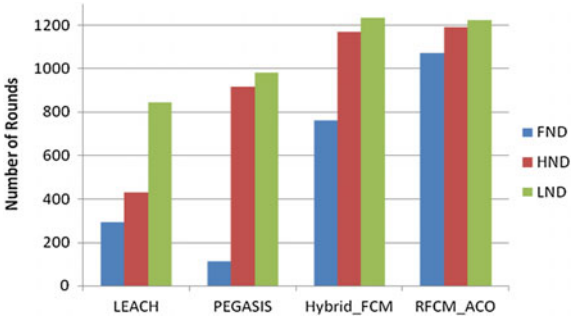


Fig. 2 Bar graph shows the FND, HND and LND in rounds for a random deployment of 100 nodes

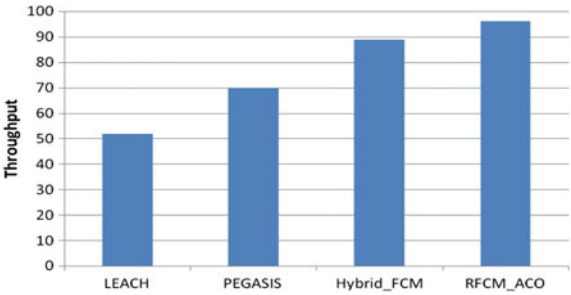


Fig. 3 Average number of data packets received at the base station per round

5 Conclusion

RFCM-ACO elongates WSN lifespan while providing a good load balance. We plan to use the other variants of FCM like possibilistic c-means (PCM) and fuzzy possibilistic c-means (FPCM) to partition the nodes. A comparative study determining the cluster compactness and validity index will be done using RFCM, PCM and FPCM. The concept of node mobility and heterogeneity is to be added for dealing with applications like robotic movement, animal tracking to name a few. The possibility of our proposal being incorporated in sensor cloud can be investigated.

Acknowledgements The authors would like to thank Prof. Utpal Biswas, Dept. of Computer Science and Engineering, University of Kalyani for his valuable suggestions. The authors would further like to thank the members of the Biomedical Imaging and Bioinformatics Lab (BIBL), Indian Statistical Institute, Kolkata for their support.

References

1. Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E.: Wireless sensor networks: a survey. Elsevier J. Comput. Netw. **38**, 393–422 (2002)
2. Yick, J., Mukherjee, B., Ghosal, D.: Wireless sensor network survey. Comput. Netw. **52**(12), 2292–2330 (2008)
3. Heinzelman, W.R., Chandrakasan, A.P., Balakrishnan, H.: An application-specific protocol architecture for wireless microsensor networks. IEEE Trans. Wireless Commun. **4**, 660–670 (2002)
4. Lindsey, S., Raghavendra, C.: Data gathering algorithm in sensor networks using energy metrics. IEEE Trans. Parallel Distrib. Syst. **13**(9), 924–935 (2002)
5. Maji, P., Pal, S.K.: RFCM: a hybrid clustering algorithm using rough and fuzzy sets. Fund. Inform. **80**, 475–496 (2007)
6. Bezdek, J.C., Ehrlich, R., Full, W.: FCM: the fuzzy c-means algorithm. J. Comput. Geosci. **10** (2–3), 191–203 (1984)
7. Dorigo, M., Sttuzle, T.: Ant Colony Optimization. MIT Press (2004)
8. Zadeh, L.A.: Fuzzy = computing with words. IEEE Trans. Fuzzy Syst. **4**(2), 103–111 (1996)

9. Hadjila, M., Guyennet, H., Feham, M.: A Hybrid Cluster and Chain Based Routing Protocol for Lifetime Improvement in WSN. Lecture Notes in Computer Science, vol. 8458. Springer International Publishing, Switzerland (2014)
10. Lam, Q.T., Hrong, M.F.: A High Energy Efficiency Approach Based on Fuzzy Clustering Topology for Long Lifetime in Wireless Sensor Network. Advanced Methods for Computational Collective Intelligence, SCI 457, pp. 367–376. Springer, Berlin (2013)
11. Chen, J.: Improving life time of wireless sensor networks by using fuzzy c-means induced clustering. In: IEEE World Automation Congress (WAC), pp. 1–4 (2012)
12. Chourasia, M.K., Panchal, M., Shrivastav, A.: Energy efficient protocol for mobile wireless sensor networks. In: Proceedings of the IEEE International Conference on Communication Control and Intelligent Systems (CCIS), pp. 79–84. IEEE (2015)
13. Nayak, P., Devulapalli, A.: A fuzzy logic-based clustering algorithm for WSN to extend the network lifetime. IEEE Sens. J. **16**(1), 137–144 (2016)
14. Sharma, T., Kumar, B.: F-MCHEL: fuzzy based master cluster head election leach protocol in wireless sensor network. Int. J. Comput. Sci. Technol. **3**(10), 8–13 (2012)
15. He, S., Dai, Y.: A clustering routing protocol for energy balance of WSN based on genetic clustering algorithm. Proc. Comput. Sci. IERI **2**, 788–793 (2012)
16. Alia, O.M.: A Decentralized Fuzzy C-Means-Based Energy-Efficient Routing Protocol for Wireless Sensor Networks, pp. 647281–647290. The Scientific World Journal, Hindawi Publishing Corporation (2014)
17. Kamal, M., Shawkat, S.A.: Two stage fuzzy logic based clustering approach wireless sensor network LEACH protocol. In: Proceedings of the IEEE International Conference on Computer and Information Technology, pp. 154–159. IEEE (2014)
18. Julie, E.G., Selvi, S.T.: Development of Energy Efficient Clustering Protocol in Wireless Sensor Network Using Neuro-Fuzzy Approach, p. 5063261. The Scientific World Journal, Hindwai Publishing Corporation. (2016)
19. Tomar, G.S., Sharma, T., Kumar, B.: Fuzzy based ant colony optimization approach for wireless sensor network. Wireless Pers. Commun. (Springer) **84**, 361–375 (2015)
20. Alami, H.E., Najid, A.: Energy efficient fuzzy logic cluster head selection in wireless senso networks. In: Proceedings of the International Conference on Information Technology for Organizations Development, pp. 1–7. IEEE (2016)

AI Doctor: An Intelligent Approach for Medical Diagnosis

Sumit Das, S. Biswas, Aditi Paul and Aritra Dey

Abstract In recent decades, artificial intelligence (AI) has found numerous applications in the field of medicine. In this paper, the key idea is that doctors are not available at emergency or doctors may do some mistakes by coincidence. To overcome or eliminate these problems, here we consider some diseases and design a prototype. In this work, we investigate different types of diseases according to various infectious agents (bacteria, virus, and parasite) and the respective diseases caused by them. Afterward, we make a formulation so that it deals with history of the disease and prescribing the proper medications according to the disease detected.

Keywords Artificial intelligence (AI) · APES · Artificial Intelligence Doctor (AIDr) · Facts · Rule · Knowledge-Base (KB)

1 Introduction

Healthcare crisis are increasing in the world especially in tropical countries due to inadequate amount of medical practitioners and poor medical facilities [1]. Due to this rate of mortality of patients especially in rural areas has increased. Consequently, artificial intelligence can help reduce this problem. Moreover, by using AI,

S. Das (✉)
Information Technology, JIS College of Engineering, Kalyani 741235, India
e-mail: sumit.it81@gmail.com

S. Biswas
DETS, University of Kalyani, Kalyani 741235, India
e-mail: biswas.su@gmail.com

A. Paul · A. Dey
Electrical Engineering, JIS College of Engineering, Kalyani 741235, India
e-mail: aditipaul40@gmail.com

A. Dey
e-mail: aritradey33@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_17

the diagnosis of patients suffering from high risk diseases can be identified instantly within a fraction of a second, by merely typing the symptoms of the patients, thereby giving an instant treatment [2]. Although AI is still in its preliminary development stage, physicians, and hospitals in the developed country are using AI techniques to identify patients who might be at a risk of kidney failure, cardiac arrest, or postoperative stress so as to prevent hospital readmission [3]. Moreover they can be made to give alerts when a particular prescribed drug can no longer be used for a patient of certain genetic type [4]. As a result of which doctors can prescribe alternative medication with a single click. It is a swift and seamless process [5]. Using PROLOG for detection of fever and its consequences and also suggesting their proper medication makes the whole process of doctor patient interaction very approachable and useful [6]. PROLOG which stands for Programming in Logic is a programming language used in artificial intelligence. It is a declarative language. In this language, different situations (facts and rules) and goals (queries) are specified and then the PROLOG interpreter derives the solution based on the situations specified [7]. It is useful in some problem solving areas like AI, natural language processing but useless in others.

2 Background Study

Decision-making in medical field is not only a complex process but it is also time-consuming. Moreover it can lead to errors in the diagnosis of a particular medical problem, which in turn will lead to incorrect medication [7]. As the relationship between symptoms and diseases are not always simple, there are many cases in which the symptoms and diseases overlap, in such situations diagnosis of a particular disease becomes difficult [8]. In the field of medicine, there are many variables which need to be considered while making decision. This can cause difference in the opinions of practitioners. So, by using an effective tool which take into consideration all the factors and also show results in difficult situations will help in solving the problems [9]. With such motivation, a small group of talented computer scientists and healthcare practitioners have developed a research program for a new discipline of study known as artificial intelligence or AI [10].

Artificial Intelligence or AI can be defined as the branch of academic field that deals with creating machines that can solve problems in human-like fashion. AI research was revived and at the beginning of twenty-first century, AI achieved great recognition worldwide and it is being used in various fields [11]. To implement problems in AI, different programming languages have been developed and used. Lisp and Prolog are the two most popular AI languages used for implementing problems in AI [12]. Prolog was first designed by Alain Colmerauer in Marseilles France around 1970. It is declarative programming language and it is mainly used for logical and reasoning problems. In this paper, AI not only deals with the diagnosis of different fever based on various infectious agents but it also deals with

prescribing proper medication [13]. As it is a very common medical illness, this system will not only help healthcare practitioners but also patients in finding proper and accurate remedies [14].

3 Methodology

A Prolog expert system (APES) supporting querying and extending the knowledge base from a command-line interface using a format oriented on natural language, with the aim of being maintainable by the domain expert (i.e., without requiring programming skills). The input format is defined as a Definite Clause Grammar (DCG), a grammar description language built into Prolog, and is therefore well defined and extensible. The most popular approach to parsing in Prolog is Definite Clause Grammar (DCG) which is a generalization of Context Free Grammar (CFG). Parsing is one of the important applications of Prolog and Logic Programming. The knowledge is represented in text files containing Prolog predicates with defined operators to make it a domain-specific language which is not too far from the actual input format, making it maintainable by the domain expert without requiring Prolog or general programming skills. This leads to the core idea and goal: an expert system where the knowledge base cannot only be queried by the domain expert, but can also be built and maintained without requiring programming skills, making this a nondomain-specific expert system that could be used to built expert systems for any domain by experts in that particular domain, not by programmers. The following Fig. 1 is showing the architecture of our work.

Fever can be defined as an abnormal rise in body temperature above the normal range. It can be caused due to a variety of medical factors like infection, cancer, and many more [15]. Mild fever without any serious medical conditions does not require any medication. Self-care measures like rest and drinking fluids may be enough [16]. However, high fever is fatal as it is seldom accompanied by serious medical conditions or illness [17]. As such it becomes important to know the agents

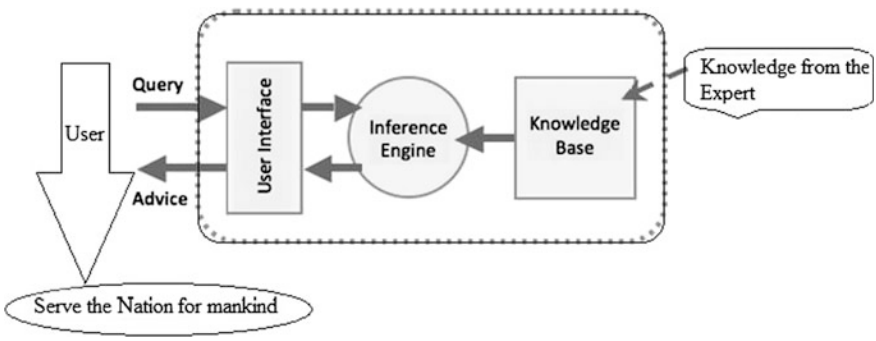


Fig. 1 Architecture of AIDr

responsible for the abnormal rise in body temperature [18]. In many cases, same symptoms arise in more than one disease which makes identifying the diseases even more difficult for a physician [19]. At such cases, AI can be used to overcome the above problem [20].

In this paper, fever has been classified according to the infectious agents (bacteria, virus, and parasite).

Fever can be classified as follows [21]:-

Bacterial fever usually accompanied by sputum, cough, and breathlessness. Some examples are: Urinary tract infection (URI) and Respiratory tract infection (RTI). RTI can be further divided into: Upper respiratory tract infection (URI) and Lower respiratory tract infection (LRI). LRI can lead to diseases like pneumonia, Bronchitis, and Viral fever—usually accompanied by rashes. Examples are: Hepatitis A, Influenza and Dengue. Parasitic fever—caused by various parasitic agents. Examples are Malaria, Typhoid, and Filarial.

3.1 Description of the Work

As respiratory tract infection (RTI) is caused both due to bacteria and virus, so in the decision tree rtib denotes RTI caused due to bacterial infection and rtiv denotes RTI caused due to viral infection. Similarly, urib denotes upper respiratory infection due to bacteria and uriv denotes upper respiratory infection due to virus. Likewise, lrib denotes lower respiratory infection due to bacteria and lriv denotes lower respiratory infection due to virus, bronchitisb denotes bronchitis due to bacteria and bronchitisv denotes bronchitis due to virus, pneumoniab denotes pneumonia due to bacteria and pneumoniav denotes pneumonia due to virus. For instance in this paper AIDr recognize the diseases like flue, back-pain based on the symptoms perceived.

Based on the above classification a decision tree is made, which is as follows.

3.2 Implementation of AI Tool

To implement the above concept the following facts and rules are applied:

FACTS: A fact must start with a predicate (which is an atom) and end with a full stop. The predicate may be followed by one or more arguments which are enclosed by parentheses. In a Prolog program, a presence of a fact indicates a statement that is true. An absence of a fact indicates a statement that is not true.

For example: ?-father(john, mary).

RULES: A rule can be viewed as an extension of a fact with added conditions that also have to be satisfied for it to be true. It consists of two parts. The first part is similar to a fact (a predicate with arguments). The second part consists of other clauses (facts or rules which are separated by commas) which must all be true for the rule itself to be true. These two parts are separated by ":-". For example, grandfather(X, Y):- father(X, Z), parent(Z, Y).

It means that "grandfather(X, Y)" is true if both "father(X, Z)" and "parent(Z, X)" are true. The comma between the two conditions can be considered as a logical-AND operator.

RECURSIVE RULE: The recursion in any language is a function that can call itself until the goal has been succeed. In Prolog, recursion appears when a predicate contain a goal that refers to itself.

In Prolog and in any language, a recursive definition always has at least two parts. A first fact that acts like a stopping condition and a rule that call itself simplified. At each level the first fact is checked. If the fact is true then the recursion end. If not the recursion continues. In this paper we will use these facts and rules to detect and classify different types of fever according to the various infectious agents along with some common diseases caused by them and their medication. Using the facts, rules, relations, symptoms, and disease-decision tree as shown in Fig. 2, a prolog program is designed and implemented by using recursive rule as shown in Fig. 3.

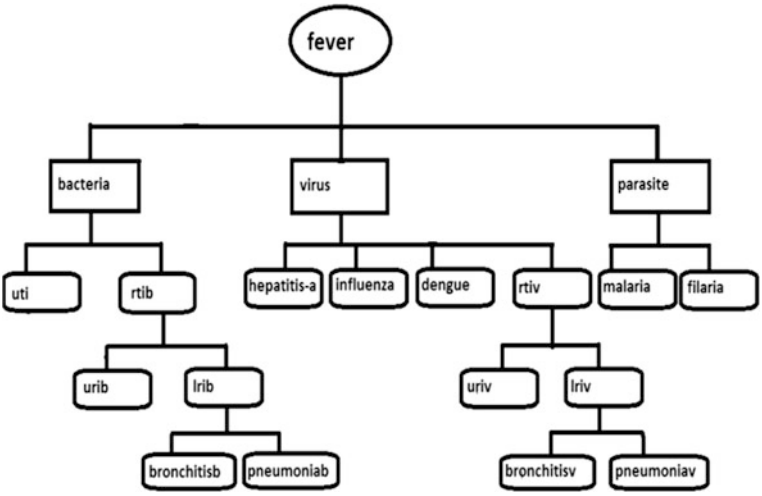


Fig. 2 Decision tree showing classification of fever based on different infectious agents and some diseases caused by them

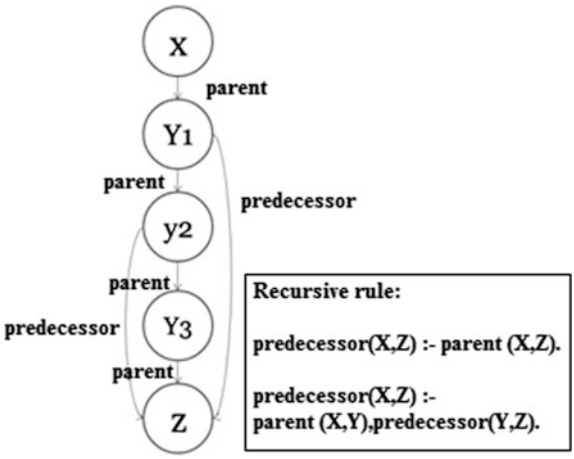


Fig. 3 Diagrammatic representation of recursive rule

4 Result and Analysis

It is to be noted that after designing the prolog program, the respective files where the program is written is consulted in the SWI-prolog environment. After that various queries are typed to get the desired result.

Now, detection of a variety of diseases can be performed by typing various queries.

For example, in the above decision tree fever is caused due to three infectious agents; now by typing a suitable query, prolog will automatically give the name of the agents along with their medication as seen in the following screenshot.

Here bactmed denotes general medication for bacterial fever which might include the use of different types of antibiotics and also includes personal hygiene and other remedies. Viralmed denotes general medication for viral fever which might include various preventive measures, some anti-viral drugs, vaccination, etc. The paramed denotes medication for parasitic fever which might include various preventive measures, use of drug as shown in Fig. 4.

Likewise some diseases caused by above-mentioned infectious agents can also be determined by typing appropriate queries in the prolog environment as shown in the following screenshots. In Fig. 5, some diseases caused dueto bacterial and viral infections are shown. By typing the query ?bacteria(X) in the prolog environment, the system automatically gives the names of the diseases caused by bacteria like in this case urinary tract infection denoted here by uti and respiratory tract infection denoted here by rtib is shown. Again respiratory tract infection can be divided into upper respiratory tract infection and lower respiratory tract infection which is denoted here by urib and lrib, respectively, which can also obtained by typing ?rtib(X) in the prolog environment. Now, lrib can lead to diseases like bronchitis and pneumonia which is denoted here by bronchitisb and pneumoniab respectively which can be obtained by typing ?lrib(X) in the prolog environment. Likewise by typing ?viral(X) some diseases caused by viral infection can be obtained as shown in Fig. 5.

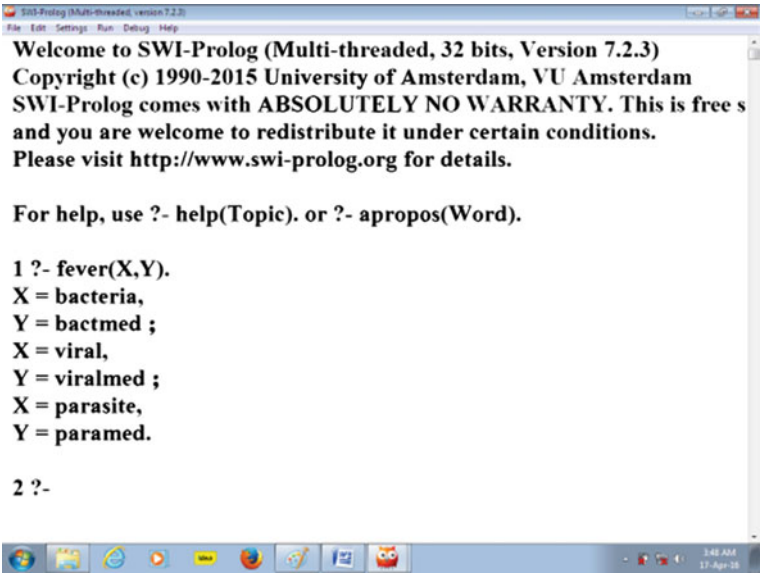


Fig. 4 Different type of fever along with their medication

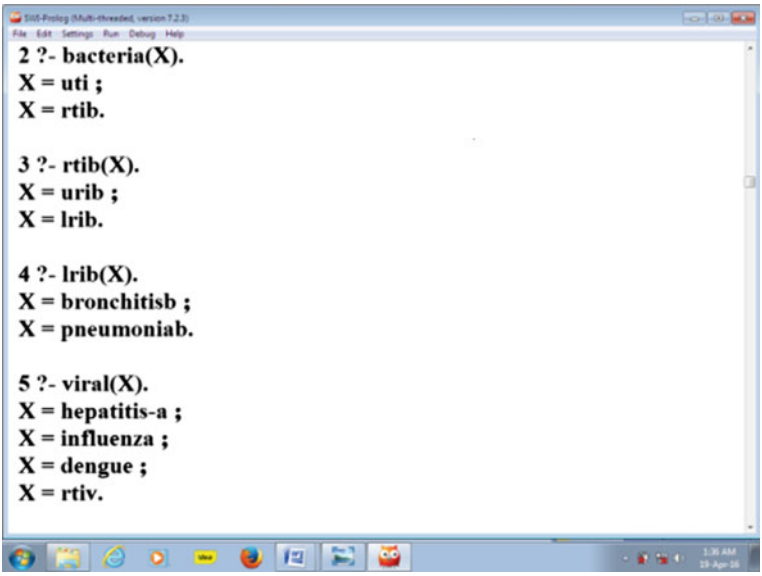


Fig. 5 Some diseases caused due to bacterial and viral infection

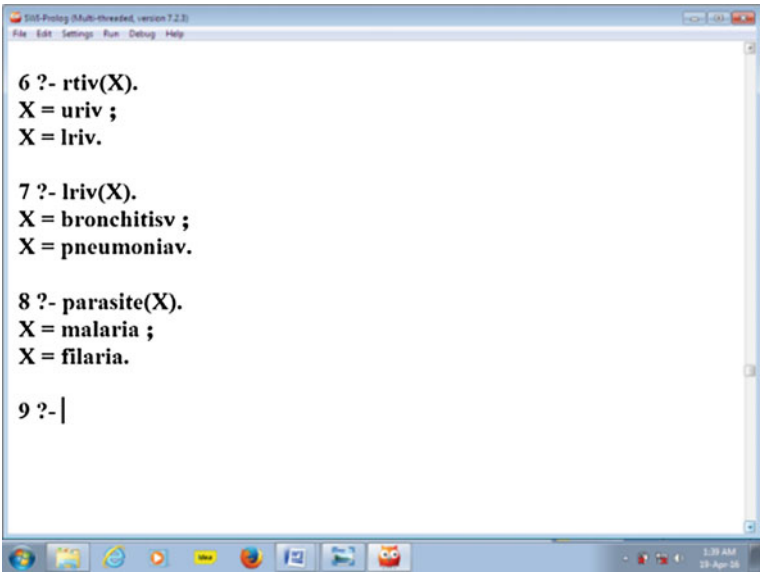


Fig. 6 Diseases caused due to parasitic infection, different types of rtiv and some disease which occur due to lriv

Similarly by typing ?parasite(X) some diseases by parasitic infection can be obtained as shown in Fig. 6. Like rtib (respiratory tract infection due to bacteria), rtiv (respiratory tract infection due to virus) can also be divided into uriv (upper respiratory tract infection due to virus), and lriv (lower respiratory tract infection due to virus) can also be obtained by typing the query ?rtiv(X), likewise, by typing ?lriv(X) some diseases which can occur due to lower respiratory tract infection can be obtained which in this case is denoted by bronchitisv (bronchitis due to viral infection) and pneumoniav (pneumonia due to viral infection).

Another characteristic of this method is that different medications for the above-mentioned diseases can also be obtained. for example, suppose someone has pneumonia, then by typing ?pneumonia(X, Y), the system gives pneumoniab (pneumonia caused due to bacteria) and pneumonia (pneumonia caused due to virus) to make it understandable to the user, the system also gives the names of the infectious agents along with the diseases, like in this case bacteria and virus are shown along with pneumoniab and pneumoniav, respectively. Now, suppose the person has pneumonia due to a viral infection then by typing the query ?pneumoniav(X), different medications for pneumonia is obtained which in this is denoted as viralmed_1, which might include various anti-viral agents like oseltamivir, peramivir, etc. and other preventive measures [22]. Like this way different diseases can be detected and also their appropriate medicine can be obtained as shown in Fig. 7.

There are more attractive features of this work, which give maximal accurate results; one of them is shown in the following Fig. 8.

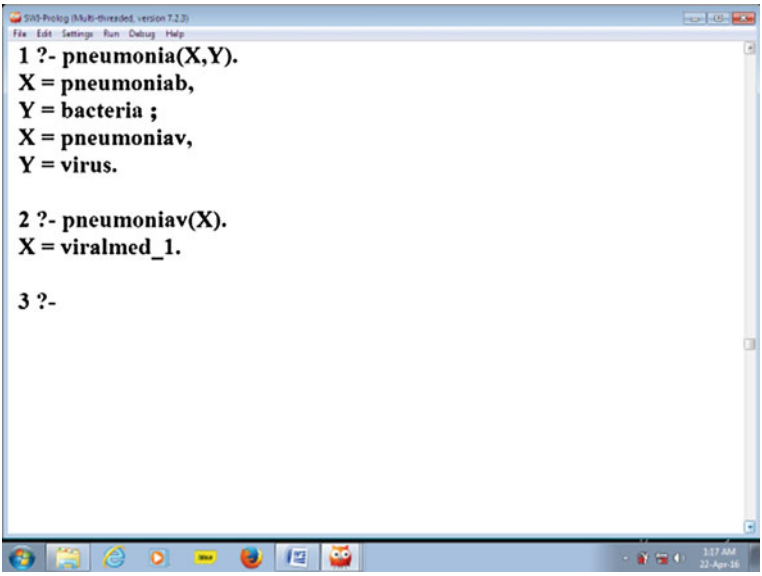


Fig. 7 Showing the system can help in detecting a disease and providing the respective medicine

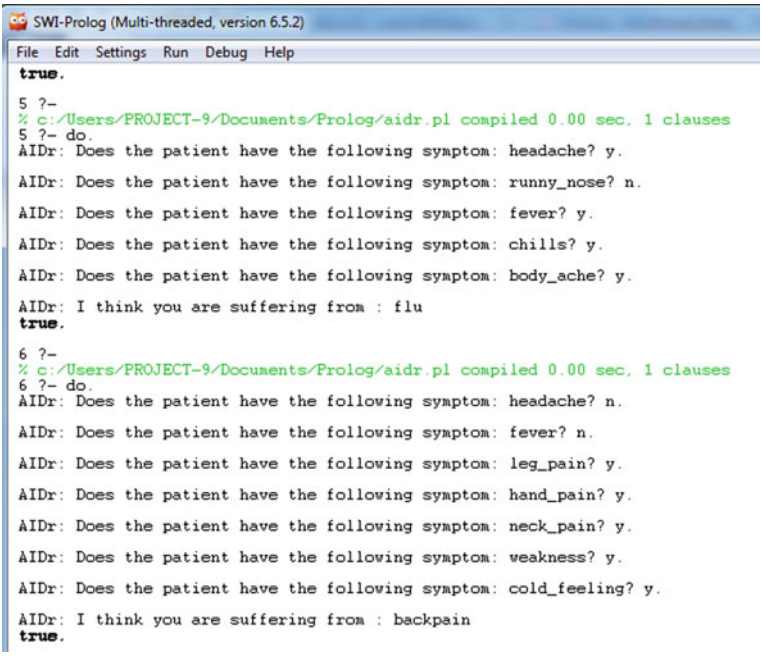


Fig. 8 Showing how AIDr asks questions and performs diagnosis

The result in Fig. 8 shows that the AIDr asks the question to the patient regarding symptoms and the patient responds accordingly. The AIDr finally responds with a proper diagnosis: the first case is "flue" and second case is "backpain."

The capability of AIDr is fully depends on the designer of the knowledge base (KB).

5 Conclusion and Future Scope

Use of AI in medical field has proved to be an effective tool in providing proper diagnosis and hence in providing proper medication. In this paper different types of disease, based on various infectious agents has been implemented using SWI-prolog which show better diagnosis compare to others [23]. It can also be used for diagnosing other diseases. Even though other methods for implementing various diseases using AI have already been studied but this method is much more user friendly as it is simple and lucid [24]. This method is still in its developing stage. With further study and assessment, medical professionals would be able to use it in fields like diagnosis, treatment of diseases, providing proper medication, development of auto-prescription generating machine and many more [25].

With further study and research assessment, we enhance the system in such way that the medical professionals would be able to use it in fields like diagnosis, treatment of diseases, providing proper medication, development of auto-prescription generating machine, and many more in sustainable manner.

References

1. Gudu, J., Gichoya, D.: A proposed diagnostic expert system for high blood pressure during pregnancy using SWI-prolog. Int. J. Artif. Intell. Mechatron.
2. Mary, A., Abdul Majeed, K.M.: Artificial intelligence and medical science: a survey. Int. J. Adv. Res. Eng. Manag. (IJAREM)
3. Sikchi, S.S., Sikchi, S., Ali, M.S.: Artificial intelligence in medical diagnosis. Int. J. Appl. Eng. Res. **7**(11) (2012). ISSN 0973-4562
4. Awwalu, J., Garba, A.G., Ghazvini, A., Atuah, R.: Int. J. Comput. Theory Eng. **7**(6) (2015)
5. Kumar, P.S., Praya, S.D., Indrajit, M.: An expert system for diagnostic of human diseases. Int. J. Comput. Appl. (0975-8887) **1**(13) (2010)
6. Haug, P.J.: Uses of diagnostic expert systems in clinical care. In: Proceedings of the Annual Symposium on Computer Application in Medical Care, pp. 379–383 (1993)
7. Grosan, C., Abraham, A.: Rule-Based Expert Systems, vol. 17, pp. 49–185. Intelligence Systems Reference Library (2011)
8. Hope, B.G., Wild, R.H.: An expert support system for service quality improvement. In: Proceedings of the Twenty-Seventh Annual Hawaii International Conference on System Science (1994)
9. Turban, E.: Expert System and Applied Artificial Intelligence. Macmillan Publishing Company, New York (1992)

10. Giarratano, J.C., Riley, G.D.: Expert System Principles and Programming, 3rd edn, pp. 624 (1998)
11. Russell, S., Norwig, P.: Artificial Intelligence and a Modern Approach, 2nd edn. Prentice Hall (2002)
12. http://www.differencebetween.com/ difference between prolog and lisp
13. Jackson, P.: Introduction to Expert Systems, 3rd edn. Addison Wesley Longman, Harlow (1999)
14. E.P. Ephzibah School of Information Technology and Engineering, VTT University, Vellore, Tamil Nadu, India: A Hybrid Genetic Fuzzy Expert System for Effective Heart Disease Diagnosis
15. Shu-Hsien, L.: Expert methodologies and application, a decade review from 1995–2004. Expert Syt. Appl. **28**, 93–103 (2005)
16. Azaab, S., Abu Naser, S., Sulisel, O.: A proposed expert systems for selecting exploratory factor and analysis procedures. J. Coll. Educ. **4**(2), 9–26 (2000)
17. Karagiannis, S., Dounis, A., Chalastras, T., Tiropanis, P., Papachristos, D.: Design of expert system for search allergy and selection of the skin tests using clips. Int. J. Inf. Technol. **3**(1) (2006)
18. Madkour, A.F., Mohammed, M.: Superguide for Diagnosis and Treatment
19. Shortliffe, E.H.: The Adolescence of AI in Medicine, 5 April 1993 (1993)
20. http://www.indianetzone.com
21. http://www.buzzle.com/articles/types-of-fever.html
22. http://emedicine.medscape.com/article/300455-medication
23. Amanda Page, Health centre: 9 ways artificial intelligence affecting medical field, 9 April (2012)
24. Kantrowitz, M.: A Timeline of Artificial Intelligence (1994)
25. Lallemand, N.C.: Health Policy Brief: Reducing Waste in Health Care. Health Affairs, 13 Dec 2012 (2012)

An Online Trend Detection Strategy for Twitter Using Mann–Kendall Non-parametric Test

Sourav Malakar, Saptarsi Goswami and Amlan Chakrabarti

Abstract Twitter is one of the most popular online social networking and micro-blogging service that enables its users to post and share text-based messages called Tweets. The data generated daily in terms of tweets are enormous and represents a rich source of information. To elicit actionable intelligence, various natural language processing (NLP) and text mining techniques are applied. Detecting of trends from twitters represents an important set of problems with a wide variety of applications and has huge appeal to diverse communities. In this paper, a simple trend detection technique based on term frequency has been proposed. In the first step, term document matrix of the tweet stream is created and top words are identified. The top word list is dynamically updated based on new streams. Time series is generated for the top words. Trends of the words are detected using Mann–Kendall non-parametric test. The method has been applied on few topical twitter datasets and proved to be quite effective.

Keywords Twitter · Text mining · Time series

1 Introduction

Undoubtedly twitter is one of the most active social networking sites as far as the number of users or activities by the user is concerned. It represents enormous opportunity of actionable intelligence. However, typically the data from social networking site is noisy, with lot of spams, abbreviations, multilingual portions, and spelling errors. Also, as it is unstructured data, statistical analysis cannot be applied

S. Malakar · S. Goswami (✉) · A. Chakrabarti
Kolkata, India
e-mail: saptarsi007@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_18

trivially on the data. Various machine learning and linguistic techniques are applied to put a structure to the data and then extract meaningful insight from the data. Twitter has been employed in variety of tasks. Some recent and interesting application includes study on unemployment [1], understanding sentiment in politics [2], understanding topics in health care [3], for psychological analysis. Trend detection is one of the important tasks in twitter. As pointed out in [1] these trends are important to news reporters, analysts, online marketing professional, and opinion tracking companies. It is also important to retailers, government entities, and more.

There are some existing trend detection tools or applications already available [4, 5]. Many of them are based on key words or busty keywords. In article [6, 7], the authors have improved trend detection using topic models which are based on latent Dirichlet allocation [8]. The trend detection techniques can be classified as [9], (i) Online and (ii) Retrospective. Current works are more focused on online detection of trends, rather than on historical corpora given the dynamic nature of social media. In this paper, we have described our approach which uses a traditional text processing techniques coupled with a strong statistical trend detection technique.

The appeal of our current work is as follows:

- It is tested against real and contemporary datasets collected using the Twitter API.
- Some portion of the corpora is built meticulously as described by the process [10], and hence is devoid of noise or spam.
- It is a lightweight technique as we use term document frequency as the basis and employ Mann-Kendal non-parametric test for identifying the trends.

The proposed method is generic in nature, so can be extended to any set text collections as example for bunch of emails, chats, reviews, etc.

The rest of the paper is organized as follows in Sect. 2; different works in trend detection are discussed Sect. 3, outlines the proposed methodology. Section 4, contains the details of experimental setup. Sections 5 and 6 contain the results and conclusion, respectively.

2 Related Work

Various studies have addressed trend detection using different techniques over Twitter in recent years. Some of the works like [4, 5] uses a keyword-based approach of trend detection, while in [6, 11], authors have proposed trend detection based on topic modeling which uses LDA. It is also observed that the emphasis is

more on detecting online trends rather than against the historical datasets. Some papers focused on geographical trend detection [12], they rely heavily on geo-coding which has practical limitations, as till today only handful devices use location awareness. In paper [13], authors have narrowed the scope of trends in a particular locality and for a particular period. Paper [14], brings forth an interesting application, where trend detection was applied after a natural disaster namely the great east Japan earthquake. In [15], authors have talked about how multiple events can propagate at the same time. They proposed a method based on grams co-occurrence and topic ranking, which was shown to be effective over a few states of the arts.

For trend detection, Mann–Kendall non-parametric test can be used. It has been used mostly for natural time series like hydrology, stream flow [16], climate change [17], and precipitation [18]. Only seldom it has been seen to be in use outside of natural time series like in finding trend in images [19] or for software age analysis [20]. As per our literature study, we could not find any efforts where Mann–Kendall test have been used for trend detection on tweets or text.

3 Proposed Methodology

In this section, the proposed methodology in terms of sequential steps has been elaborated. First of all details of preprocessing steps have been discussed, subsequently the entire flow is represented using a diagram and finally the pseudo codes have been outlined.

Data Collection: Tweets from varied topics have been collected using the Twitter API. While some of the datasets have been created for business news and poll, some of the tweets have been collected for socially sensitive topics. Details of the datasets have been enclosed in Sect. 4.

Preprocessing: Slandered techniques of Text Mining like punctuation removal, URL removal, stop word removal, and stemming have been performed. Domain specific stop words are also removed.

Term Frequency Generation: A tweet set corresponding to a particular time period (Window) has been taken as the input. A tweet set can be generated from a particular user's timeline; it can also be done based on keyword search. Different preprocessing steps as described earlier are applied and term document matrix (tdm) is generated. Based on the tdm, top k terms are selected. For the next window, again top k terms are selected and these lists are merged using union. The process is repeated for all available windows and also will be applicable to online mode.

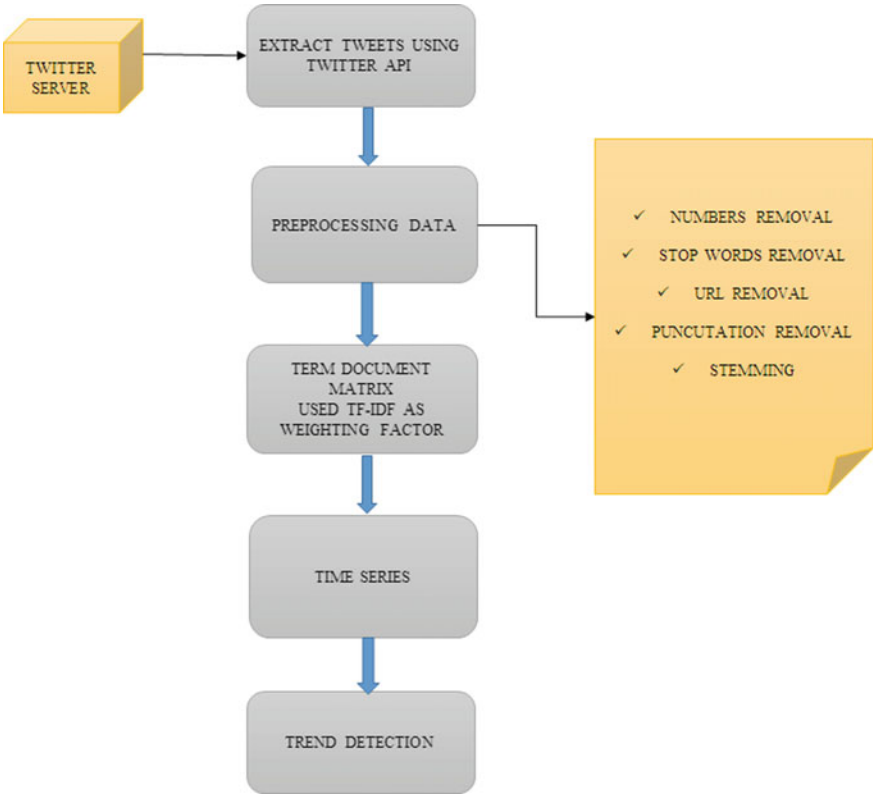


Fig. 1 The diagrammatic representation of preprocessed methodology

Trend Analysis: In this step, a time series is constructed for each of the words as obtained from earlier test. Based on the Mann–Kendall non-parametric test, trend analysis is performed. If the p-value is only less than a threshold (α), then it is taken as a statistically significant trend. Based on the sign of 's,' the direction of the trend is determined.

The diagrammatic representation is shown in Fig. 1. It is followed by the algorithm representation of the methodology.

```
Method: TopWordExtractor
   Input: Text documents belonging to a window W, K
   Output: Top K words and their frequency
   Begin
         Step 1: Apply pre-processing like stop word removal, Stemming, removal of
      punctuation, number etc.
         Step 2: Prepare term document
         Step 3: Top K words are selected with their frequency.
   End

Method: TrendDetector
   Input: Multiple stream (S) of text documents with window D, α
   Output: Trends ( TW, Sign) As a two dimensional array
   Begin
   GTW<- {}
   TW<-{}
   Index < -0
   Trends<-{}
   For each S
       TW <-TopWordExtractor(S)
       GTW <−GTW ∪TW
       For each distinct w1 ∈ GTW
        Create a time series (tsw1) of w1
        For each time series, perform Mann-Kendal test for trends to calculate s and the
corresponding p-value.
       If p < α
          Index< -index +1
          Trend [Index, 1] <- w1
          If s < 0
                Trend [Index, 2] <- -1
          else
                Trend [Index, 2] <- 1
       End if
   End For
End
```

4 Materials and Methods

For our experimental setup, we have collected total four sets of tweets according to different domain, such as West Bengal Election, 2016, Abesh incident, 2016, JNU incident, 2016 and some business tweets according to Economic Times in the year 2014 using Twitter API and where R [21] is used as a computational environment, where 'twitteR' [22] package is used. For each data set, we have applied some Text Mining techniques like stemming [23], stop word removal [23], punctuation and URL removal [12] for preprocessing the data. The different datasets that have been used are summarized in Table 1.

For text mining package 'tm' [23] and for Time Series package 'zoo' [24] has been used, respectively. For statistical significance testing, p-value of 0.2 have been taken. The value of k for selecting top keywords has been taken as 50.

Table 1 Details of Twitter Datasets

Dataset name	Extraction method	Number of tweets	Time period	Window size
Economic Times	From user timeline	10,602	July–November 2014	15 days
Bengal Election	By keywords	9340	April–May 2016	4 days
JNU incident	By keywords	132,000	2016/02/19 to 2016/02/29	1 day
Abesh incident	By keywords	958	2016/07/25 to 2016/08/2	1 day

5 Results and Discussion

We have analyzed the results from two perspectives. In the first perspective, we have shown that what is the total time required for processing each data set to demonstrate the suitability of the method to be adopted in an online scenario and in the second subsection we have discussed about trending topics and their significance.

5.1 Time Taken in Preprocessing and Trend Detection

In the below table, the average processing times in minutes have been reported. It can be observed from Table 2.

- The average time taken to extract the top keywords taken on an average 1–15 min.
- Average time to detect trend is even smaller as compared to time taken to process window.
- Hence the above prescribed method is quite suitable for online detection of trends.

Table 2 Dataset with total processing time

Data set	Average time taken to process window	Average time to detect trends (s)
Economic Times	2.99 min	2
Bengal Election	0.95 min	1
JNU incident	15 min	5
Abesh incident	3 s	0.60

5.2 Trends and Their Significance

In this section, the details of trend found from the top words are discussed. The results are presented in Table 3.

It can be seen, that for all of the datasets some meaningful trends could be found, some of which if known earlier can help different organizations, policymakers to take timely decisions.

The time series of 'alliance' and 'azadi' are shown in Figs. 2 and 3, respectively.

Table 3 Summary of trends

Data set	Trends	s-value	p-value	Conclusion
Economic Times	Railbudget	−8	0.17524	Rail budget started as a trend, but faded away, may be because it did not meet expectation
Bengal Election	Alliance	−5	0.14856	The alliance showed clear negative trend in Bengal Election, much before the results were out
JNU incident	Azadi	24	0.036278	Azadi remained a theme with positive trend throughout the JNU incident
	Protest	−17	0.061116	Protests slowly ran out of stream
Abesh incident	Murder	−13	0.20841	The incident was about demise of a 17 year old. The trends show, how the general focus on murder got removed

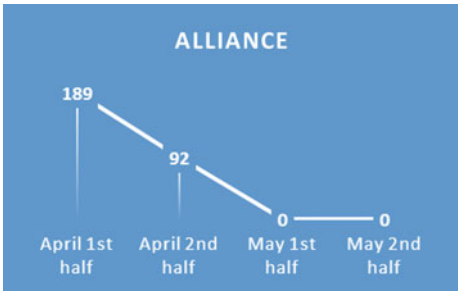


Fig. 2 Time series of 'alliance'

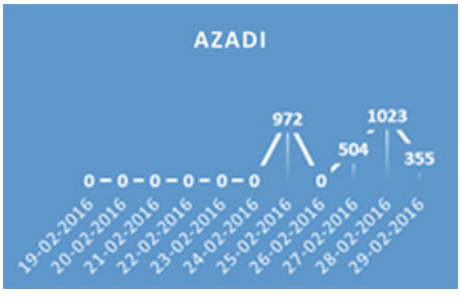


Fig. 3 Time series of 'azadi'

6 Conclusions

Twitter represents a largely untapped source of actionable intelligence. Trend detection on tweets, can unearth important information which is valuable to diverse communities. In this paper, a lightweight trend detection technique has been proposed. To put the data in structured format, typical text processing techniques have been used followed by preparation of term document matrix. Based on a window, top keywords are extracted and time series has been constructed with a dynamic vocabulary update. Mann–Kendall non-parametric test has been used for detection of trends. The technique has been applied on contemporary, real-time tweet datasets collected by the authors, employing either keyword-based or user timeline-based extraction. The above method seems to be quite effective in detecting trends. It may be noted, though the techniques have been applied on retrospective datasets, the method taken quite less time in both preprocessing as well as trend detection and hence is well suited for online detection of trends as well. As a future work, the tuning of various parameters to suit a particular tweet stream will be adopted.

References

1. Nirmala, C.R., Roopa, G.M., Naveen Kumar, K.R.: Twitter data analysis for unemployment crisis. In: 2015 International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), pp. 420–423. IEEE (2015)
2. Singhal, K., Agrawal, B., Mittal, N.: Modeling Indian general elections: sentiment analysis of political Twitter data. In: Information Systems Design and Intelligent Applications, pp. 469–477. Springer India (2015)
3. Barnwal, A.K., Choudhary, G.K., Swarnim, R., Kedia, A., Goswami, S., Das, A.: Application of Twitter in health care sector for India. In: 3rd IEEE International Conference on Recent Advances in Information Technology (Accepted)
4. Mathioudakis, M., Koudas, N.: Twitter monitor: trend detection over the twitter stream. In: Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data, pp. 1155–1158. ACM (2010)
5. Kim, D., Kim, D., Rho, S., Hwang, E.: Detecting trend and bursty keywords using characteristics of Twitter stream data. Int. J. Smart Home **7**(1), 209–220 (2013)
6. Lau, J.H., Collier, N., Baldwin, T.: On-line trend analysis with topic models: Twitter trends detection topic model online. In: COLING, pp. 1519–1534 (2012)
7. Bolelli, L., Ertekin, Ş., Giles, C.L.: Topic and trend detection in text collections using latent Dirichlet allocation. In: European Conference on Information Retrieval, 6 April 2009, pp. 776–780. Springer, Berlin
8. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet allocation. J. Mach. Learn. Res. **3**(Jan), 993–1022 (2003)
9. Naaman, M., Becker, H., Gravano, L.: Hip and trendy: characterizing emerging trends on Twitter. J. Am. Soc. Inf. Sci. Technol. **62**(5), 902–918 (2011)
10. Kumar, S., Maskara, S., Chandak, N., Goswami, S.: Article: empirical study of relationship between Twitter mood and stock market from an Indian context. Int. J. Appl. **8**, 33–37

11. Bolelli, L., Ertekin, S., Lee Giles, C.: Topic and trend detection in text collections using latent Dirichlet allocation. In: ECIR 09 Proceedings of the 31st European Conference on IR Research on Advances in Information Retrieval, pp. 776–780
12. Walther, M., Kaisser, M.: Geo-spatial event detection in the twitter stream. In: Advances in Information Retrieval, pp. 356–367. Springer, Berlin (2013)
13. Ishikawa, S., Arakawa, Y., Tagashira, S., Fukuda, A.: Hot topic detection in local areas using Twitter and Wikipedia. In: ARCS Workshops (ARCS), 2012, pp. 1–5. IEEE (2012)
14. Miyabe, M., Miura, A., Aramaki, E.: Use trend analysis of Twitter after the Great East Japan earthquake. In: Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work Companion, pp. 175–178. ACM (2012)
15. Aiello, L.M., Petkos, G., Martin, C., Corney, D., Papadopoulos, S., Skraba, R., Goker, A., Kompatsiaris, I., Jaimes, A.: Sensing trending topics in Twitter. IEEE Trans. Multimed. **15**(6), 1268–1282 (2013)
16. Yue, S., Pilon, P., Cavadias, G.: Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. J. Hydrol. **259**(1), 254–271 (2002)
17. Guhathakurta, P., Sreejith, O.P., Menon, P.A.: Impact of climate change on extreme rainfall events and flood risk in India. J. Earth Syst. Sci. **120**(3), 359–373 (2011)
18. Liu, Q., Yang, Z., Cui, B.: Spatial and temporal variability of annual precipitation during 1961–2006 in Yellow River Basin, China. J. Hydrol. **361**(3), 330–338 (2008)
19. Neeti, N., Ronald Eastman, J.: A contextual Mann-Kendall approach for the assessment of trend significance in image time series. Trans. GIS **15**(5), 599–611 (2011)
20. Grottke, M., Li, L., Vaidyanathan, K., Trivedi, K.S.: Analysis of software aging in a web server. IEEE Trans. Reliab. **55**(3), 411–420 (2006)
21. R Core Team: R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org/ (2013)
22. Gentry, J.: Twitter: R based Twitter Client. R package version 1.1.9. https://CRAN.R-project.org/package=twitteR (2015)
23. Feinerer, I., Hornik, K., Meyer, D.: Text mining infrastructure in R. J. Stat. Softw. (2008)
24. Zeileis, A., Grothendieck, G.: Zoo: S3 infrastructure for regular and irregular time series. J. Stat. Softw. **14**(6), 1–27 (2005). http://jstatsoft.org/v14/i06/

Part IV
Digital Signal and Image Processing

Full Reference Image Quality Assessment: A Survey

Baisakhi Sur Phadikar, Goutam Kumar Maity and Amit Phadikar

Abstract This article presents a brief review on full-reference (FR) image quality assessment (IQA) techniques. The discussion starts with traditional Peak Signal-to-Noise Ratio (PSNR) and then gradually moves to complex IQA models based on wavelets (or its various variants) and human visual system (HVS). The techniques are discussed from mathematical perspective to theoretical viewpoint. These discussions will be very useful for relevant researchers to have an apparent understanding about the status of recent FR-IQA. Few research problems are also discussed as an outcome of the article.

Keywords IQA · Full-reference IQA · UQI · SSIM · VIF · GMSD · FSIM

1 Introduction

The emergent needs for digital imaging technologies in applications like, biomedical systems, medical imaging, and communications have highlighted the need for exact quality evaluation methods. Many processes can change the quality of images. Those are compression, transmission, and acquisition. Therefore, perfect assessment of the image quality is a significant step in several image-based applications. The aim of objective IQA is to design an algorithm that can

B.S. Phadikar (✉)
Department of Computer Science & Engineering, MCKV Institute of Engineering, Liluah, India
e-mail: baisakhi.sp@gmail.com

G.K. Maity
Department of ECE, Netaji Subhash Engineering College, Garia, India
e-mail: goutam123_2005@yahoo.co.in

A. Phadikar
Department of Information Technology, MCKV Institute of Engineering, Liluah, India
e-mail: amitphadikar@rediffmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_19

automatically assess the quality of images in a perceptually steady manner. According to the accessibility of an original image, objective image quality metrics can be classified as full-reference, no-reference, or reduced-reference.

In the compact representation many state-of-the-art IQA (especially FR algorithms) take up a general two-stage arrangement. In the first phase, image distortion/quality is calculated locally. The locality may be defined in scale, space, and orientation. To translate such quality maps into a single quality value, a pooling stage is used in the second phase of IQA algorithm. The present article gives a very concise analysis of related research works on FR-IQA. Moreover, few open research problems are also highlighted as an outcome of this article. The prediction performance and computation time of these methods are also highlighted in this article.

The organization of the article is as follows: Sect. 2 describes related works for FR-QIA in brief. Conclusion and few open research problems are highlighted in Sect. 3.

2 FR-IQA Indices

In this section, we will discuss various FR-IQA along with their pros and cons.

2.1 PSNR/MSE (Mean Squared Error)

PSNR and MSE are the mostly used technique to measure the image quality. Those techniques rely on mathematical assessment between an original image and a distorted version of the original image. Although being exceptional to compute the image quality, those techniques do not reflect the HVS perceived quality.

2.2 Information Content Weighted (ICW) MSE and PSNR

Wang and Li [1] proposed information content weighted MSE (IW-MSE) and information content weighted PSNR (IW-PSNR) to get better performance over MSE and PSNR. Assume, $x_{j,i}$ and $y_{j,i}$ are the *i*th Laplacian pyramid transform (LPT) coefficients at the *j*th scale. The symbol $w_{j,i}$ is the information content weight computed at the corresponding location. Then IW-MSE is defined as:

$$\mathrm{IW}-MSE=\prod_{j=1}^{M}\left[\frac{\sum_i w_{j,i}\left(x_{j,i}-y_{j,i}\right)}{\sum_i w_{j,i}}\right]^{\beta_j}, \tag{1}$$

where the symbol M is the number of scales. The symbol β_j is the weight set to the jth scale. IW-MSE can be transformed to IW-PSNR as:

$$\mathrm{IW-PSNR} = 10\log_{10}\left(\frac{L^2}{\mathrm{IW}-MSE}\right). \tag{2}$$

2.3 Universal Quality Index (UQI)

UQI is easy to compute and independent of individual observation and viewing environment [2]. Let, $\mathbf{x} = \{\mathrm{x_i}|\mathrm{i} = 1, 2, \ldots \mathrm{N}\}$ and $\mathbf{y} = \{\mathrm{y_i}|\mathrm{i} = 1, 2, \ldots \mathrm{N}\}$ are the original image signal and the distorted image signal, respectively. The UQI is described as:

$$Q = \frac{4\sigma_{xy}\bar{x}\bar{y}}{(\sigma_x^2 + \sigma_y^2)\left[(\bar{x})^2 + (\bar{y})^2\right]}, \tag{3}$$

where

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \quad \text{and} \quad \bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i \tag{4}$$

$$\sigma_x^2 = \left(\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2\right) \quad \text{and} \quad \sigma_y^2 = \left(\frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{y})^2\right) \tag{5}$$

The values of Q are varied from -1 to 1. The value '1' is attained when $\mathbf{y} = \mathbf{x}$ and '-1' is happened when $\mathrm{y_i} = 2\bar{\mathrm{x}} - \mathrm{x_i}$ for all i = 1, 2 … N.

2.4 Structural Similarity Index Measure (SSIM)

The SSIM scheme divides the job of similarity calculation into three comparisons, i.e., (a) luminance, (b) contrast, and (c) structure [3]. First of all, the luminance of each signal (i.e., **x** and **y**) is evaluated. This is anticipated as the average intensity, i.e.,

$$\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i. \tag{6}$$

Second, the system eliminates the average intensity from the signal. The resulting signal $(\mathbf{x} - \mu_x)$, in discrete form, corresponds to the projection of vector '**x**' onto the hyper plane described as:

$$\sum_{i=1}^{N} x_i = 0. \tag{7}$$

The system then calculates the signal contrast based on the standard deviation. This is defined as:

$$\sigma_x = \left(\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu_x)^2 \right)^{1/2}. \tag{8}$$

Similarly, the contrast comparison, i.e., $c(\boldsymbol{x}, \boldsymbol{y})$ is the function of σ_x and σ_y. Third, both the signals are normalized (divided) through its own standard deviation (σ_x and σ_y). The structure comparison, i.e., $s(\boldsymbol{x}, \boldsymbol{y})$ is conducted on these normalized signals $(x - \mu_x)/\sigma_x$ and $(y - \mu_y)/\sigma_y$. Finally, those three elements are joined together to yield a final similarity measure,

$$S(x, y) = f(l(x, y), c(x, y), s(x, y)) \tag{9}$$

It is to be noted that the three components are relatively independent. Luminance comparison $l(x, y)$ is calculated by:

$$l(x, y) = \frac{2\mu_x \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}. \tag{10}$$

Constant C_1 is included to avoid instability when $\mu_x^2 + \mu_x^2$ is very close to zero. The contrast comparison is defined as:

$$c(x, y) = \frac{2\sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}. \tag{11}$$

Similarly, structure comparison is defined as:

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3}, \tag{12}$$

where σ_{xy} is calculated as:

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu_x)(y_i - \mu_y). \tag{13}$$

Finally, luminance, contrast, and structure comparisons of (10)–(12) are combined together to estimate the SSIM index.

$$SSIM(x, y) = [(l(x, y)]^{\alpha} . [c(x, y)]^{\beta} . [s(x, y)]^{\gamma}, \tag{14}$$

where $\alpha > 0$, $\beta > 0$, and $\gamma > 0$. The SSIM index is expressed as:

$$SSIM(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_x^2 + C_2)}. \tag{15}$$

2.5 Multi-scale SSIM (MS-SSIM)

Multi-scale technique is an efficient way to include image details at different resolutions [4]. The scheme repeatedly applies a low-pass filter on the reference and distorted image signals and then down samples the filtered image by a factor of 2. The system index the original image signal as scale 1. The highest scale is scale M, which is achieved after M – 1, iterations. Then the contrast comparison (Eq. 11) and the structure comparison (Eq. 12) at *j*th scale are measured and those are denoted as c_j(x, y) and s_j(x, y), respectively. The luminance comparison (Eq. 10) is calculated only at *M* scale. This is denoted as l_M(x, y). The final MS-SSIM is calculated by combining the measurement of different scales as:

$$\mathrm{MS-SSIM}(\mathbf{x}, \mathbf{y}) = [\mathrm{l_M}(\mathbf{x}, \mathbf{y})]^{\alpha_M} \cdot \prod_{\mathrm{j}=1}^{\mathrm{M}} [\mathrm{c_j}(\mathbf{x}, \mathbf{y})]^{\beta_j} [\mathrm{s_j}(\mathbf{x}, \mathbf{y})]^{\gamma_j}. \tag{16}$$

To adjust the relative weight of different components in (16) the exponents, i.e., α_M, β_j and γ_j are used.

2.6 Information Content Weighted (ICW) MS-SSIM

Let, $\mathbf{x_{j,i}}$ and $\mathbf{y_{j,i}}$ be the *i*th local image patches at *j*th scale. Assume, the number of assessment windows in the above scale is $\mathrm{N_j}$. Then SSIM at *j*th scale is described as:

$$\mathrm{SSIM_j} = \frac{1}{N_j} \sum_i c(x_{j,i}, y_{j,i}) s(x_{j,i}, y_{j,i}), \tag{17}$$

for $j = 1, \ldots, M - 1$ and

$$\mathrm{SSIM_j} = \frac{1}{N_j} \sum_i l(x_{j,i}, y_{j,i}) c(x_{j,i}, y_{j,i}) s(x_{j,i}, y_{j,i}), \tag{18}$$

for $j = M$.

The MS-SSIM is given by:

$$\mathrm{MS-SSIM} = \prod_{j=1}^{M} \left(\mathrm{SSIM_j}\right)^{\beta_j}, \tag{19}$$

Afterwards, information content weighted SSIM (IW-SSIM) is calculated by merging ICW with MS-SSIM [1]. Assume, $w_{j,i}$ is the information content weight calculated at the *i*th spatial position in the *j*th scale. Then the *j*th scale IW-SSIM is given as:

$$\mathrm{IW-SSIM_j} = \frac{\sum_i w_{j,i} c\left(x_{j,i}, y_{j,i}\right) s\left(x_{j,i}, y_{j,i}\right)}{\sum w_{j,i}}, \tag{20}$$

for $j = 1, \ldots, M - 1$ and

$$\mathrm{IW-SSIM_j} = \frac{1}{N_j} \sum_i l(x_{j,i}, y_{j,i}) c\left(x_{j,i}, y_{j,i}\right) s\left(x_{j,i}, y_{j,i}\right), \tag{21}$$

for $j = M$. The finishing overall IW-SSIM is defined by:

$$\mathrm{IW-SSIM} = \prod_{j=1}^{M} \left(\mathrm{IW-SSIM_j}\right)^{\beta_j}. \tag{22}$$

The scheme has higher computational complicity than the MSSIM as it uses LPT.

2.7 Visual Information Fidelity (VIF) Criterion

Main theoretical problem with SSIM is the assumption about the structural information [5]. On the other hand, VIF assumes that HVS has evolved to be the best perceive for natural scenes. The scheme models the original image as an outcome of a natural source. The mutual information between the output of the source and the HVS channel is the quantity of information that brain can extract.

Natural source is represented as Gaussian scale mixtures (GSM). This is a random field (RF). It can be described as a product of two independent RFs, i.e., $\mathcal{C} = \mathcal{S} \cdot \mathcal{U} = \{\mathcal{S}_i \cdot \vec{\mathcal{U}}_i : i \in I\}$. The symbol '$\boldsymbol{I}$' is a set of spatial indices. $\mathcal{S} = \{\mathcal{S}_i : i \in I\}$ is an RF of positive scalars. $\mathcal{U} = \{\mathcal{U}_i : i \in I\}$ is a Gaussian vector RF with zero mean and covariance $\mathbf{C}_\mathrm{U}$. The subbands of wavelet domain are divided into nonoverlapping blocks of M coefficients. Then the block 'i' is expressed as M-dimensional vector $\vec{C}_i$.

Distortion channel is expressed as a signal deviation and additive noise. The model of the output of the distortion channel looks like $\mathcal{D} = \mathcal{G}.\mathcal{C} + \mathcal{V} = \left\{g_i \vec{C}_i + \vec{V}_i : i \in I\right\}$. The RF $\mathcal{G} = \{g_i : i \in I\}$ is a deterministic scalar

gain field. The symbol $\mathcal{V} = \left\{ \vec{V}_i : i \in I \right\}$ is an additive, stationary, zero mean, white Gaussian noise with variance $C_V = \sigma_v^2 I$.

HVS channel is modeled as a single additive noise component that adds uncertainty to the signal that flows through the HVS. It is stationary and has zero mean, white Gaussian noise $\mathcal{N} = \left\{ \vec{N}_i : i \in I \right\}$ for an original and $\mathcal{N}'$ for a test image. Then $\mathcal{E} = \mathcal{C} + \mathcal{N}$ and $\mathcal{F} = \mathcal{D} + \mathcal{N}'$ are the signals processed by the brain for an original and a test image, respectively.

From this model, the mutual information for both images $I\left(\vec{C}, \vec{F} \middle| s\right)$ and $I\left(\vec{C}, \vec{E} \middle| s\right)$ can be computed, where 's' stands for the realization of '*S*' for particular reference image. Final measure then combines these results into the signal formula:

$$VIF = \frac{\sum_{j \in subbands} I\left(\vec{C}^{N,j}, \vec{F}^{N,j} | s^{N,j}\right)}{\sum_{j \in subbands} I\left(\vec{C}^{N,j}, \vec{E}^{N,j} | s^{N,j}\right)}, \tag{23}$$

where $\vec{C}^{N,j}$ represents the N elements of RF C_j for jth subband.

(*A*) *Interpretation of VIF-Based upon Information Content Weighting*: The VIF scheme [5] does not look to be fit into the two-stage structure. This is due to the fact that the information content is computed over the whole image prior to the fidelity ratio is calculated.

$$VIF = \frac{\sum_i I(R_i; F_i | s_i)}{\sum_i I(R_i; E_i | s_i)}. \tag{24}$$

However, with some simple alteration, VIF can be understood by the two-stage structure. Then, VIF can be described as:

$$VIF = \frac{\sum_i w_i \mathrm{VIF_i}}{\sum_i w_i}, \tag{25}$$

where the local VIF evaluation is defined as:

$$VIF_i = \frac{I(R_i; F_i | s_i)}{I(R_i; E_i | s_i)}, \tag{26}$$

and a weighting function

$$w_i = I(R_i; E_i | s_i). \tag{27}$$

2.8 Gradient Magnitude Similarity Deviation (GMSD) Index

The gradients of image are sensitive to image degradation. This motivates authors in [6] to use gradient-based local quality map for overall image quality measurement.

(*A*) *Gradient Magnitude Similarity* (*GMS*): For digital images, the gradient is evaluated by convolving an image signal with operators (is a kind of liner filter) viz. classic Roberts, Scharr, Sobel, Prewitt, or some application specific ones [6]. The Prewitt operator for vertical (*y*) and horizontal (*x*) directions are described by:

$$\mathbf{h}_x = \begin{bmatrix} 1/3 & 0 & -1/3 \\ 1/3 & 0 & -1/3 \\ 1/3 & 0 & -1/3 \end{bmatrix}, \ \mathbf{h}_y = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 \\ -1/3 & -1/3 & -1/3 \end{bmatrix}. \tag{28}$$

The horizontal and vertical gradient of images **r** and **d** are calculate by convolving $\mathbf{h}_x$ and $\mathbf{h}_y$ with the reference (**r**) and distorted images (**d**). The gradient magnitudes at location '*i*,' indicated as $\mathbf{m}_r(i)$ and $\mathbf{m}_d(i)$. Those are calculates as:

$$m_r(i) = \sqrt{(r \otimes h_x)^2(i) + \left(r \otimes h_y\right)^2(i)}, \text{ and } m_d(i) = \sqrt{(d \otimes h_x)^2(i) + \left(r \otimes h_y\right)^2(i)} \tag{29}$$

where the symbol '⊗' represents the convolution action. GMS is computed as:

$$GMS(i) = \frac{2m_r(i)m_d(i) + c}{m_r^2(i) + m_d^2(i) + c}. \tag{30}$$

The positive constant '*c*' provides the numerical stability. It is also to be noted that $\mathbf{m}_r(i)$ or $\mathbf{m}_d(i)$ is computed from a small local area. Obviously, if $\mathbf{m}_r(i)$ and $\mathbf{m}_d(i)$ are equal, *GMS*(*i*) will attain the maximal value, i.e., one.

(*B*) *Pooling with Standard Deviation*: The final quality value of an image can be projected from the GSM via a pooling phase. The average pooling which is the most commonly used pooling strategy is described as:

$$GMSM = \frac{1}{N} \sum_{i=1}^{N} GMS(i), \tag{31}$$

where '*N*' is the number of pixels in the image. Then authors propose to compute gradient magnitude similarity deviation (GMSD) as:

$$\text{GMSD} = \sqrt{\frac{1}{\text{N}} \sum_{i=1}^{\text{N}} (\text{GMS(i)} - \text{GMSM})^2}. \tag{32}$$

Note that for the high GMSD value, the signal distortion is large.

2.9 Feature Similarity Index Measure (FSIM)

To evaluate the quality of an image, FSIM makes use of two kinds of information, i.e., phase congruency (PC) and image gradient magnitude (GM).

(A) *Phase Congruency* (*PC*): The PC [7] model assumes that Fourier coefficients with highest phase hold the points where image features are perceived by HVS. At first, a set of response vectors are calculated at location '**x**,' orientation 'o,' and scale 's' using 2-D log-Gabor filter bank for the reference and test images. Then, the local amplitude and energy of these vectors are calculated. Lastly, the value of PC at location '**x**' is evaluated by:

$$\mathrm{PC}(\mathrm{x}) = \frac{\sum_{\mathrm{o}} \mathrm{E}_{\mathrm{o}}(\mathrm{x})}{\varepsilon + \sum_{\mathrm{s}} \sum_{\mathrm{o}} \mathrm{A}_{\mathrm{s,o}}(\mathrm{x})}, \tag{33}$$

where symbol $\mathrm{A}_{\mathrm{s,o}}(\mathrm{x})$ is the local amplitude and $\mathrm{E}_{\mathrm{o}}(\mathrm{x})$ is the local energy. PC(x) is a real number. It takes its values in the interval [0, 1].

(B) *Gradient Magnitude* (*GM*): Gradient magnitude (GM) of an image f(**x**) is described as $\mathrm{G} = \sqrt{\mathrm{G}_{\mathrm{x}}^2 + \mathrm{G}_{\mathrm{y}}^2}$. The symbols $\mathrm{G}_{\mathrm{x}}(\mathbf{x})$ and $\mathrm{G}_{\mathrm{y}}(\mathbf{x})$ represent the partial derivatives of the image f(**x**) along horizontal and vertical directions by any gradient operators (liner filter), respectively.

(C) *Feature Similarity Index Measure* (*FSIM*): The system splits the feature similarity measurement of $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ into two mechanisms, each for PC or GM. At the first phase, the similarity measure for $PC_1(\mathbf{x})$ and $PC_2(\mathbf{x})$ is calculated as:

$$S_{PC}(x) = \frac{2PC_1(x) \cdot PC_2(x) + T_1}{PC_1^2(x) \cdot PC_2^2(x) + T_1}, \tag{34}$$

where the constant 'T_1' is introduced to amplify the stability of S_{PC}. In the same way, the GM values, i.e., $G_1(\mathbf{x})$ and $G_2(\mathbf{x})$ are compared. Similarity measure is defined as:

$$S_G(x) = \frac{2G_1(x) \cdot G_2(x) + T_2}{G_1^2(x) \cdot G_2^2(x) + T_2}, \tag{35}$$

Similarly, 'T_2' is a constant. Then, $S_{PC}(\mathbf{x})$ and $S_G(\mathbf{x})$ are pooled to get the overall similarity $S_L(\mathbf{x})$ for $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$. $S_L(\mathbf{x})$ is calculated by:

$$S_L(x) = [S_{PC}(x)]^{\alpha}[S_G(x)]^{\beta}, \tag{36}$$

where the symbol 'α' and 'β' are used to regulate the relative significance of PC and GM, respectively. For a given location '**x**,' if either $f_1(\mathbf{x})$ or $f_2(\mathbf{x})$ has a major PC value, it reflects that the location '**x**' offers higher impact on HVS and in the calculation of similarity between f_1 and f_2. Therefore, authors set $PC_m(\mathbf{x}) = \max(PC_1(\mathbf{x}), PC_2(\mathbf{x}))$ to give the significance of $S_L(\mathbf{x})$ in the overall similarity measurement and is defined as:

$$FSIM = \frac{\sum_{x\in\Omega} S_L(x)PC_m(x)}{\sum_{x\in\Omega} PC_m(x)}, \tag{37}$$

where, Ω means the whole image spatial domain.

2.10 Color Image Quality Assessment

The first color image quality measure is proposed in [8]. In this scheme, a simple model of human color vision is explained which quantitatively describes different perceptual factors, e.g., brightness and saturation. The FSIM can be extended to color image also. First of all, the red-green-blue (RGB) color images are transformed into YIQ color space. In YIQ color space, the symbol Y stands for the luminance information and the symbols I and Q express the chrominance information. Let, Q_1 (Q_2) and I_1 (I_2) be the Q and I chromatic channels of the image f_1 (f_2), respectively. Similar to the description of $S_G(\mathbf{x})$ and $S_{PC}(\mathbf{x})$ as discussed above, the similarity between chromatic features is expressed as:

$$S_I(x) = \frac{2I_1(x)\,I_2(x) + T_3}{I_1^2(x) + I_2^2(x) + T_3}, \quad S_Q(x) = \frac{2Q_1(x)\,Q_2(x) + T_4}{Q_1^2(x) + Q_2^2(x) + T_4}, \tag{38}$$

where the symbols T_3 and T_4 are the positive constants and are used to add stability to the system as discussed in SSIM. Then $S_I(\mathbf{x})$ and $S_Q(\mathbf{x})$ are merged to get the chrominance similarity measure and is denoted by $S_C(\mathbf{x})$, for $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ as:

$$S_C(x) = S_I(x) \cdot S_Q(x). \tag{39}$$

Finally, FSIM is expanded to $FSIM_C$ by incorporating chromatic information and is expressed by:

$$FSIM_c = \frac{\sum_{\mathbf{x}\in\Omega} S_L(x) \cdot [S_C(x)]^{\lambda} \cdot PC_m(x)}{\sum_{\mathbf{x}\in\Omega} PC_m(x)}, \tag{40}$$

where parameter ($\lambda > 0$) is used to regulate the significance of chromatic information.

Table 1 Comparative performance of various competing FR-IQA models on the LIVE image databases. *Source* [6]

FR-IQA model	LIVE (779 image signals)			Execution time (s)
	SRC	PCC	RMSE	
PSNR	0.8760	0.8720	13.360	0.0016
SSIM	0.9480	0.9450	08.950	0.0388
VIF	**0.9640**	**0.9600**	**07.610**	1.1745
MS-SSIM	0.9510	0.9490	8.6190	0.1379
GMSM	0.9600	0.9560	8.0490	0.0079
IW-SSIM	0.9570	0.9520	08.350	0.5196
FSIM	**0.9630**	**0.9600**	07.670	0.5269
GMSD	0.9600	**0.9600**	**07.620**	0.0110

Table 1 shows the comparative performance of various competing FR-IQA models in terms of Pearsons correlation coefficient (PCC), Spearman rank coefficient (SRC), and root-mean-square error (RMSE) on the LIVE databases.

3 Conclusions and Scope of Future Works

Image quality assessment is a fast rising field of research in recent years. The algorithms that are highlighted are growing rapidly. Of course, only a few numbers of methods are reported in the paper. The reported algorithms are extensively cited in the literature and also have high-quality performance [9]. Another reason for the selection is that the source code has been made available online for the most of these methods. So the concerned readers can execute them and also regenerate the reported results.

From the review, it is cleared that the entire schemes extract image features to design an IQA. However, no work found till now in literature which address content-based image retrieval (CBIR) using various IQA. In other words, more specifically, how to use image information (features) that are use for FR image quality measurement to be used for CBIR. In CBIR systems input provide in terms of an image and based on image attribute (features) matching, the most similar images from the database are retrieved. By considering all these issues, one can make the correct selection for each specific application like CBIR.

References

1. Wang, Z., Li, Q.: Information content weighting for perceptual image quality assessment. IEEE Trans. Image Process. **20**(5), 1185–1198 (2011)
2. Wang, Z., Bovik, A.C.: A universal image quality index. IEEE Signal Process. Lett. **9**(3), 81–84 (2002)
3. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)

4. Wang, Z., Simoncelli, E.P., Bovik, A.C.: Multi-scale structural similarity for image quality assessment. In: Proceedings of 37th IEEE Asilomar Conference on Signals, Systems and Computers, Pacific Grove, pp. 1398–1402 (2003)
5. Sheikh, H.R., Bovik, A.C.: Image information and visual quality. IEEE Trans. Image Process. **15**(2), 430–444 (2006)
6. Xue, W., Zhang, L., Mou, X., Bovik, A.C.: Gradient magnitude similarity deviation: a highly efficient perceptual image quality index. IEEE Trans. Image Process. **23**(2), 684–695 (2014)
7. Zhang, L., Zhang, L., Mou, X., Zhang, D.: FSIM: a feature similarity index for image quality assessment. IEEE Trans. Image Process. **20**(8), 2378–2386 (2011)
8. Faugeras, O.D.: Digital color image processing within the framework of a human visual model. IEEE Trans. Acoust. Speech Signal Process. **27**, 380–393 (1979)
9. Liu, A., Lin, W., Narwaria, M.: Image quality assessment based on gradient similarity. IEEE Trans. Image Process. **21**(4), 1500–1512 (2012)

A Framework for Face Recognition Based on Fuzzy Minimal Structure Oscillation and Independent Component Analysis

Sharmistha Bhattacharya (Halder) and Srijita Barman Roy

Abstract This paper aims to provide a better accuracy for face recognition procedure. This new algorithm is based on accurate feature extraction and proper classification. In this paper feature coordinate-based ICA is used for feature extraction. Pixel values of invariable coordinates (containing decisive data) for every training set are considered for analyzing through ICA. After feature extraction, these values are used for fuzzy minimal structure oscillation-based classification. Proposed face recognition procedure accentuates improved classification considering the feature vectors, which is the outcome of independent component analysis of the face image.

Keywords Fuzzy minimal structure oscillation · Feature coordinate · Feature coordinate-based ICA · Face recognition

1 Introduction

In general topological space, the concept of minimal or m_x^* structure was presented by Popa and Noiri, in 2000 [1]. The vital concept of $-V$ set and $-\Lambda$ set was first introduced by Maki [2] in 1986. But in Fuzzy environment Alimohammady and M. Roohi first introduced the concept of fuzzy minimal structure [3]. A fuzzy set oscillates between at least two fuzzy open set and two fuzzy closed set. It was first introduced by Mukherjee and Halder in 2007 [4]. Fuzzy m_x^* Oscillation was first presented by Bhattacharya (Halder) and Roy in 2010 [5]. Image classification

S. Bhattacharya (Halder) (✉) · S.B. Roy
Department of Mathematics, Tripura University, Agartala, India
e-mail: halder_731@rediffmail.com

S.B. Roy
e-mail: Srijita.barmanroy@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_20

algorithm based on Fuzzy m_x^* Oscillation was introduced by us in [6], where we show a successful application of Fuzzy m_x^* Oscillation in the field of face recognition.

Based on different numerical properties of images, different classification algorithms could be found in existing literatures. The influential properties of classifiers are based on either any statistical process or any distribution free process to extract class description. So we are motivated to extract independent components rather we can say, the basis images from some different facial image frame. Linear transformation of the basis images form the perfect m_x^* structure. We introduce of a new algorithm for face recognition and show its functionality for recognizing an unknown image [7] from a set of known images. In this paper, the proposed algorithm was implemented using Matlab7.9 and verified on various available Face images collected from FacePix database [8] considering a set of important feature values.

2 Review of Literature

Different mathematical and statistical methods have been proposed to introduce better face recognition procedure. Papers [9–15] highlight some ICA-based face recognition algorithm. In [10] authors partitioned the facial image in many parts, and perform modular ICA. With nearer neighbour classification face recognition task is performed. The paper [12] author give stress on feature selection which is based on Genetic algorithm. After this GA-based ICA, new features are classified by nearest algorithm. All these literature use different ICA-based processing for feature extraction. Where our proposed algorithm differs from these excising literature. But feature-based minimal structure formation provides very efficient classification. In [15] author proposed a very new approach based on lattice ICA and very good learning procedure for classification. In this paper, author proposes a very simple ICA-based face recognition algorithm which classifies the localized feature of expression through Euclidian distance classification [9]. Paper [10] provides different types of image processing such land covering information images. Karimi and Zadeh [11] proposed a new classification-based sparse representation. Here they apply ICA for dimension reduction. In [13] this paper proposed method is based on extraction of independent components from edges of face images. Paper [14] provides 2d PCA and ICA-based face recognition. These are treated as features for face recognitions. Recent research literatures hardly consist of ICA-based feature extraction along with fuzzy-based classifier. To illustrate the concept of Fuzzy minimal structure oscillation, some theorems and definitions are included in the below.

2.1 Definition 1 [5]

$$\Lambda_{a_j}(x) = \inf\left\{\mu_{a_j}(x_i) : \mu_{a_j}(x_i) \geq \mu_{a_j}(x), x_i \in G \text{ is an open set, j } = 1, 2, 3 \ldots \text{n}\right\} = \widehat{I},$$
if no such open set exist. (1)

$$Int_{a_j}(x) = \sup\left\{\mu_{a_j}(x_i) : \mu_{a_j}(x_i) \leq \mu_{a_j}(x), x_i \in G \text{ is an open set, j} = 1, 2, 3 \ldots \text{n}\right\} = \phi,$$
if no such open set exist. (2)

2.2 Definition 2 [5]

An operator $O^o : I^x \to I^x$ such that $O^o a_j(x) = \Lambda_{a_j}(x) - Int_{a_j}(x)$ is said to be fuzzy m_x^*-open oscillatory operator and an operator $O^c : I^x \to I^x$ such that $O^o a_j(x) = Cl_{a_j}(x) - V_{a_j}(x)$ is said to be fuzzy m_x^* closed oscillatory operator.

2.3 Fuzzy m_x^*-Structure or Minimal Structure Oscillation on Image Comparison: [5]

In the paper [6], we introduce the fuzzy m_x^* oscillation base image classification. Suppose the database contains image $I = I_1, I_2 \ldots$ To compare an unknown image with the known image by the help of fuzzy minimal structure-oscillation, the following cases might appear according to the values of oscillatory operators.

Case 1: Now if possible let $O^o_{I_k}(\text{y}_{ij}) = 1$. $\Lambda_{I_k}(\text{y}_{ij}) = 1$ and $\text{Int}_{I_k}(\text{y}_{ij}) = 0$ i.e. the intensity of the unknown image at pixel(i, j) is not within a known position.

Case 2: For $0 < O^o_{I_k}(\text{y}_{ij}) < 1$, (i) If $O^o_{I_k}(\text{y}_{ij}) \leq 0.1$, there may be some similarity between the two images at pixel (i, j). (ii) If $O^o_{I_k}(\text{y}_{ij}) > 0.1$, Then if difference between intensity of the unknown image pixel(i, j) and $\Lambda_{I_k}(\text{y}_{ij})$ or $\text{Int}_{I_k}(\text{y}_{ij}) \leq 0.1$ then also the images may be similar at the pixel(i, j).

Case 3: $O^o_{I_k}(\text{y}_{ij}) = \Lambda_{I_k}(\text{y}_{ij}) - \phi \leftrightarrow$ If the difference between intensity of the unknown image pixel(i, j) and $\Lambda_{I_k}(\text{y}_{ij}) \leq 0.1$ then, the images may be considered as similar at the pixel (i, j). Otherwise different.

Case 4: $O^o_{I_k}(\text{y}_{ij}) = \widehat{\text{I}} - \text{Int}_{I_k}(\text{y}_{ij})$, as case 3 where infimum value is not present

3 Feature Coordinate-Based ICA

We propose here a new face recognition method using modified ICA and Fuzzy minimal structure oscillation-based classification. $X = (x_1, x_2, \ldots x_k)$ is a set where x_k is m $\times$ n matrix, which we call feature coordinate frame. x_k matrix contains feature coordinate pixel values. $x_{k_{xy}}$ specifies n observations based on first location xy of the feature coordinate. $X_{k_{xy}} = A \cdot S_{k_{xy}}$ Where, A is unknown matrix of size $m \times n$ matrix where $S = (s_1, s_2, \ldots s_k)$. Here $S_{k_{xy}}$ specifies set of ICs for x_k facial image at the location frame xy. Feature coordinate frame is an $p \times q$ coordinate frame or matrix which contains some decisive pixel values. By changing the starting location of the feature coordinate frame position can be changed. The frame size can also be modified automatically. Therefore f contains q, number of p dimension column vector. By applying Fast-ICA [16] algorithm on f pixels column vectors it will generate some independent components which is the basis images for each training set. We are inclined to use ICA, due to two main fundamental properties of ICA. Dimension reduction and feature extraction are two very useful properties of ICA. Whitening and PCA is pre-processing steps of ICA. PCA step is performed on image data before applying ICA. This step is very useful to reduce high-dimension image data, but it is not a mandatory step of ICA. We follow paper [17].

So by applying ICA we are getting the basis images which carries important feature with reduced dimension. These independent components are need to be transformed so that it is applicable to the fuzzy m_x^* structure. Suppose the set of generated Independent Components $\alpha_i = |\alpha_i| / \max_i |\alpha_i|$. Let $I_{(i,j)}$ be the (i,j)th pixel of image I. We consider a square matrix as $I_{(i+\varepsilon, j+\varepsilon)}$ $\varepsilon \geq 0$, is an integer where $p = = q$. Since the ICA signals have the values which does not lies in $[0, 1]$. So applying the Linear transform we convert the ICA signals to fuzzy membership value. Obviously $1 \in$ the new membership values. Hence it forms an m_x^* structure.

Architecture

In this frame-based ICA architecture we assign a multiple frames on face images. These frames are located at the different places which contains decisive pixel values. We read the frames pixel values of an single individual from different location (suppose eye lid, Lips corners, etc.). After that applying the ICA we get a set of basis images. Suppose $\lambda_1, \lambda_2, \ldots \lambda_n$ are ICs generated from $f_1, f_2, \ldots f_n$ of image I_j. Where dimension of λ_k is depends on the frame size i.e. q and the mixing matrix w. For ICA algorithm converges fast in ν steps where $\nu < p$, the size of the basis image can be defined as $\nu \times q$. So if kth frame's IC is

$$\lambda_k = [\nu \times q] \tag{3}$$

For the IC of frame of image I_j is,

$$f_k = [1 \times \lambda_k] \tag{4}$$

where L_j is extracted features of image I_j,

$$L_j = \sum_{k=1}^{n} \lambda_k \tag{5}$$

For this architecture m_x^* structure is defined as

$$m \times structure = \left[L; L_{j=1:m}\right] \tag{6}$$

where L is the combinations of ICs extracted from $\boldsymbol{m}$ training images.

4 Face Recognition Based on Fuzzy Minimal Structure Oscillation and Independent Component Analysis

In our proposed algorithm, at first we form the image training set from the extracted feature from each independent component. The coordinate frame of (i,j)th pixel of image I. Specifying the coordinate values and the size the variable size fame can be situated any where in the face. Depending upon the environment that means of nature of face database, the number of frame can be increased or decreased. But in the last phase of the paper we show that frame with large number of decisive pixel value the algorithm provide better result.

Algorithm: Face Recognition based on Fuzzy Minimal Structure Oscillation and Independent Component Analysis.

Input: *Frame position-based Image data extracted from any Image database, Considered as known images for training set.*

Output: *Decision if any unknown image is SIMILAR or not with the known image.*

1. *Input any image database*
2. *Set the frame size,* $(p, q) \leftarrow$ *values., Place the frame by assigning the value of* ε, i, j *in* $I_{(i+\varepsilon, j+\varepsilon)}$ *where* $I_{(i,j)}$ *is* (i,j)th *coordinate in image I,* $\varepsilon \geq 0$*, integer value, Set the location of feature coordinates. Setting the frames in the coordinates with decisive coordinate values.*
3. $fcIC_i[\nu, q] \leftarrow$ *frame-based Fast-ICA will be done per image frame basis.*
4. $fcIC_{LT}[\nu, q] = |fcIC_i[\nu, q]|/|\max(fcIC_i[\nu, q])|$ *Apply Linear Transformation on the basis images of each frame. For number of frame* $= n$
5. *For m number of training set minimal structure open set* $\rightarrow fcIC_{LT}[m, n]$
6. Input the unknown image

7. $fcICUN_{LT}[\nu, q] \leftarrow$ *Retrieve the feature value from that frame*
   Apply Fuzzy m_x^* Oscillation on $fcIC_{LT}[m, n]$ and $fcICUN_{LT}[\nu, q]$
8. *count* 0 *to q* Calculate $\Lambda\Lambda[m,n] \leftarrow MinMax.Val(fcIC_{LT}[m,n] \circ fcICUN_{LT}[count,n])$
9. Calculate *Int* $Int[m,n] \leftarrow MaxMin\_Val(fcIC_{LT}[m,n] \square fcICUN_{LT}[count,n])$
10. Calculate Open_Oscillatory_Operator

$$O^o[count, n] = \Lambda[count, n] - Int[count, n]$$

11. $if\,(count \leq q)$ Goto 10
12. Compute the Negative or Complementary image set. Calculate CL_OPTR, V_OPTR and $O^{cl}[count, n]$
13. Consider four cases of O^o and O^{cl} for pixel by pixel comparison. Decision about Similarity and Difference.

5 Experiment

We perform our experiment on FacePix database which is a face image database created at the Center for Cognitive Ubiquitous Computing (CUbiC) at Arizona State University, and it is available free of charge to the worldwide. FacePix database consist of different pose of an individuals. Face image's representations vary from angle $+90°$ to $-90°$, where $+90°$ represents a left profile view, $0°$ represents a frontal view, and $-90°$ represents a right profile view. We use Matlab7.9 for our experiment. We place frame on different places (Suppose left, right eye ball, lips, etc., as shown in Fig. 1) on the face. To test our algorithm we develop a Matlab program. If the frame is of size 6×4, then it contains 24 pixel. This pixel may carry important features of a particular face. Now applying ICA we can reduce

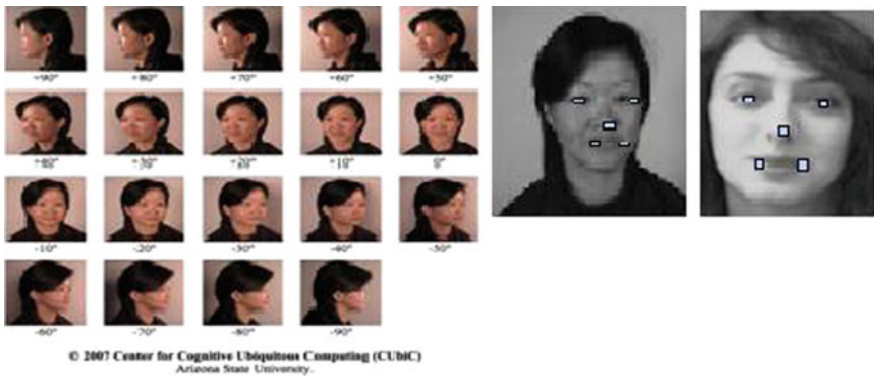


Fig. 1 (*Left*) FacePix database (used as to form minimal structure). (*Right*) Frames in training set and unknown face. Frames for ICA and minimal structure formation

Table 1 Compared with unknown image of different individual (f = 3)

Frame size		Reduced dimension		Number of Individual	Number of frame = 3 $[1, \Lambda]$	Number of similar pixel(s)	Number of different pixel(s)		Accuracy
$p \times q$	Number of pixel	$l \times k$	Number of reduced pixels				O^o	O^o value no case	
4 × 6	24	3 × 4	12	6	36	5	30	1	0.83
4 × 6	24	3 × 4	12	4	36	5	29	2	0.80
4 × 5	20	2 × 5	10	6	30	6	24	0	0.80
4 × 5	20	2 × 5	10	4	30	7	23	0	0.76
6 × 3	18	2 × 2	4	6	18	4	11	3	0.61
6 × 3	18	2 × 2	4	4	18	8	10	0	0.55

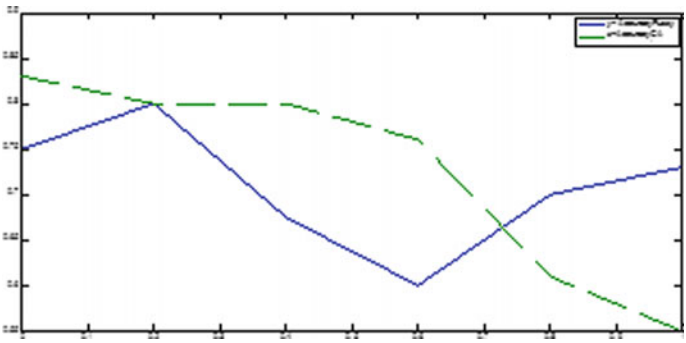


Fig. 2 Accuracy between Fuzzy m_x^* oscillation [6] (*blue*) and frame-ICA-based fuzzy m_x^* oscillation (*green dotted*). With smaller frame size data reduction and performance reduces

the size into 3 × 4 size matrix. In this case, by applying ICA, the frame size is almost 50% reduced. Now, analyzing the experimental result we induced that ICA-based feature extraction extract the vital independent components of a facial image, because of which the fuzzy minimal structure consist of less number of data. So less number of value decreases the number of comparison. Only 52% of data forms the minimal structure (Table 1, Fig. 2).

6 Conclusion

In this paper we show ICA and fuzzy minimal structure oscillation-based face recognition system. Previously we introduce this new classification algorithm with directly assembled data from face image. But number of comparison was very high as we are comparing each pixel of unknown image sets with the set of pixel values in the training image set. In this paper as we are using ICA on each frame, so number of pixel values in each training set is less. In experimental result we show ICA-based feature extraction along with dimension reduction produces very good result.

References

1. Popa, V., Noiri, T.: On M-continuous functions. AnalUnivDunarea Jos-Galati, Ser Mat Fiz MecTeorFasc II **18**(23), 31–41 (2000)
2. Maki, H.: Generalized-sets and the associated closure operator. Special issue on commemoration of Prof KazusadaIkedas Retirement, pp. 139–146 (1986)
3. Alimohammady, M., Roohi, M.: Fuzzy minimal structure and fuzzy minimal vector spaces. Chaos, Solitons Fractals **27**(3), 599–605 (2006)
4. Mukherjee, A., Halder, S.: Some more results on oscillatory region of a fuzzy set. In: International Conference on Modeling and Simulation CIT, Coimbatore, pp. 665–670 (2007)

5. Bhattacharya (Halder), S., Roy, S.: On Fuzzy m_x^* oscillation and it's application in image processing. Ann. Fuzzy Math. Inform. **7**(2), 319–329 (2014)
6. Bhattacharya (Halder), S., Barman Roy, S., Saha, S.: Application of fuzzy oscillation in the field of face recognition. In: International Symposium on Advanced Computing and Communication, IEEE, Silchar, pp. 192–107 (2015)
7. http://www.scielo.org.co/scielo.php?pid=S0120-56092013000100010&script=sci_arttext
8. FacePix-DataBase https://cubic.asu.edu/content/facepix-database
9. Lihong, Z., Ye, W., Hongfeng, T.: Face recognition based on independent component analysis. In: Chinese Control and Decision Conference (CCDC), pp. 426–429. IEEE (2011). ISSN: 1948-9439
10. Falco, N., Benediktsson, J., Bruzzone, L.: Spectral and spatial classification of hyperspectral images based on ICA and reduced morphological attribute profiles. IEEE Trans. Geosci. Remote Sens. **53**(11), 6223–6240 (2015). doi:10.1109/TGRS.2015.2436335
11. Karimi, M.M., Zadeh, S.H.: Face recognition: a sparse representation-based classification using independent component analysis. In: 6th International Symposium on Telecommunications (IST'2012), pp. 1170–1174. IEEE. 978-1-4673-2073-3
12. Yi-qiong, X., Bi-cheng, L., Bo, W.: Face recognition by fast independent component analysis and genetic algorithm. In: Proceedings of the Fourth International Conference on Computer and Information Technology (CIT'04), pp. 194–198. IEEE (2004). doi:10.1109/CIT.2004.1357196
13. Karande, J.K., Talbar, S.N.: Independent component analysis of edge information for face recognition. In: Springer Briefs In Applied Sciences And Technology, Computational Intelligence. Springer, New Delhi. ISBN 978-81-322-1512-7. doi:10.1007/978-81-322-1512-7
14. Zhang, X., Zhang, X., Ren, X.: Two dimensional principal component analysis based independent component analysis for face recognition. In: International Conference on Multimedia Technology (ICMT), pp. 934–936. IEEE, Hangzhou (2011). doi:10.1109/ICMT.2011.6002199
15. Marques, I., Gran˜a, M.: Face recognition with lattice independent component analysis and extreme learning machines. Soft Comput. **16**, 1525–1537 (2012). doi:10.1007/s00500-012-0826-4,1525-1537
16. Hyvärinen, A.: Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans. Neural Netw. **10**(3), 626–634 (1999)
17. Bartlett, M.S., Movellan, J.R., Sejnowski, T.J.: Face recognition by independent component analysis. IEEE, Proc. Trans. Neural Netw. **13**, 1450–1464 (2002)

Part V
Modern Instrumentation, Control, Robotics and Automation

Computing Reflectance of Three-Layer Surface Plasmon-Based Sensor at Visible Spectra

Pratibha Verma, Arpan Deyasi and Pratiti Paul

Abstract Critical performance of three-layer surface plasmon-based sensor is computed at visible spectra. Reflectance is calculated at wider range of incidence angle and also for different thicknesses of gold layer. A sharp dip is observed at reflectance profile when propagation constant of incident wave and that of plasmon surface becomes nearly equal. Field enhancement is also obtained for different structural parameters. Findings will play important role in designing optical sensor at visible range for bio applications.

Keywords Reflectance · Angle of incidence · Layer thickness · Field enhancement · Visible spectra

1 Introduction

Surface Plasmon is the rapid oscillation of free electrons at the interface of metal–dielectric interface [1] when dielectric constant of the metal surface is a complex quantity, and electrons are delocalized and coherent. According to Maxwell's equation, it is the progressively propagating electron density wave occurred at metal–dielectric interface [2]. It is widely used nowadays for design of biological, chemical, and molecular sensors [3–5], high-resolution spectroscopy [6], measurement at sub-nanometer level [7], etc. It exhibits resonance when propagation

P. Verma (✉) · P. Paul
Department of Electronics and Communication Engineering, NIT Agartala, Jirania, India
e-mail: vermapratibha1007@gmail.com

P. Paul
e-mail: paulpratiti@gmail.com

A. Deyasi
Department of Electronics and Communication Engineering, RCC Institute of Information Technology, Kolkata, India
e-mail: deyasi_arpan@yahoo.co.in

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_21

constant of the surface plasmon wave becomes numerically equal to that of incident surface wave [8]. This characteristic is manifested by a sharp fall in the reflectance spectra.

Ishimaru et al. [8] proposed surface plasmon resonance sensors by using metamaterials designed at microwave frequency. Charbonneau et al. [9] showed theoretical prediction for gratings and mentioned that measured optical spectra are presented for a family of third-order uniform step-in-width Bragg gratings operating in the long-range surface Plasmon-polariton mode supported by thin finite-width Au films in a homogeneous dielectric background. Popov et al. [10] showed that propagation features of the signal wave are different in NIM compared with its counterpart in natural crystals and the possibilities of compensation of strong losses by wavelength-selective frequency-tunable narrowband filtering. Ding et al. [6] and his coworkers explored the mechanism of nanoscale regions with a strongly enhanced local electromagnetic field that allow trace-molecule detection, bio-molecule analysis, and surface characterization of various materials. Sreekanth et al. [11] developed a miniaturized plasmonic biosensor platform based on a hyperbolic metamaterial. Liu et al. [12] demonstrated a fiber optic surface plasmon resonance (SPR) biosensor based on smartphone platforms. Demetriadou [13] showed that plasmonic imaging is crucial for understanding cellular behaviors for biological sciences, where is used to image and track organelles in cells, such as DNA and virus molecules. Zavats et al. [14] showed various surface plasmon modes in the optical properties of metallic nanostructures and surface plasmon-polariton behavior on periodically nano-structured metal surfaces and thin films. Satija et al. [15] demonstrated the study of the effect of dendrimeric and nondendrimeric support in surface Plasmon resonance (SPR) sensors for proteinaceous antigen analysis. Pechprasarn and Somekh [16] showed quantitative measurement of refractive index over localized regions in bio-imaging applications so that antibody-antigen binding can be measured on smaller scale.

In this paper, impact of thickness of gold layer, incident angle on reflectance and wavelength of optical source is analytically computed for three-layer SPR-based sensor. As well as impact of incidence angle on transmittance, reflectance, and field enhancement for different values of is analytically computed for both the cases. Wavelength of He–Ne laser light is taken in range of wavelength of visible range. Results will play critical role in determining SPR-based optical sensor performance.

2 Mathematical Formulation

The reflectance of perpendicular and parallel polarized waves for complex index of refraction are given by

$$r_{\perp} = \frac{n_1 \cos(\theta_{incident}) - n_2 \cos(\theta_{refracted})}{n_1 \cos(\theta_{incident}) + n_2 \cos(\theta_{refracted})} \tag{1}$$

$$r_{||} = \frac{n_2 \cos(\theta_{incident}) - n_1 \cos(\theta_{refracted})}{n_2 \cos(\theta_{incident}) + n_1 \cos(\theta_{refracted})} \tag{2}$$

Since perpendicular polarized or transverse electric light cannot excite surface plasmon resonance because of the oscillations of the electric field does not couple with the longitudinal modes of vibration of the sample media, so we have to consider parallel electric field. Fresnel coefficient for parallel polarized incident wave is

$$r_{||} = \frac{\left(\frac{n_2}{n_1}\right)^2 \cos(\theta) - i\sqrt{\sin^2\theta - \left(\frac{n_2}{n_1}\right)^2}}{\left(\frac{n_2}{n_1}\right)^2 \cos(\theta) + i\sqrt{\sin^2\theta - \left(\frac{n_2}{n_1}\right)^2}} \tag{3}$$

For physical system

$$R_{||} = \left[\frac{n_1\sqrt{1 - \left[\frac{n_1}{n_2}\sin(\theta_{incidence})\right]^2} - n_2\cos(\theta_{incidence})}{n_1\sqrt{1 - \left[\frac{n_1}{n_2}\sin(\theta_{incidence})\right]^2} + n_2\cos(\theta_{incidence})}\right]^2 \tag{4}$$

When $n_1 < n_2$, external reflection results and when $n_1 > n_2$, internal reflection occurs. If $n_1 = n_2$, then no reflection occurs because there is no optical boundary to interact with.

For multilayer, reflectance coefficient between layer i and j is

$$r_{ij} = \frac{n_j^2\sqrt{n_i^2 - n_{ambient}^2\sin^2\theta} - n_i^2\sqrt{n_j^2 - n_{ambient}^2\sin^2\theta}}{n_j^2\sqrt{n_i^2 - n_{ambient}^2\sin^2\theta} + n_i^2\sqrt{n_j^2 - n_{ambient}^2\sin^2\theta}} \tag{5}$$

For parallel polarized systems, the characteristic matrix describing each layer is

$$M_i = \begin{bmatrix} \cos\left(\frac{2\pi}{\lambda}n_i h_i \cos\theta_i\right) & -i\,\sin\left(\frac{2\pi}{\lambda}n_i h_i \cos\theta_i\right) \\ -i\sqrt{\frac{\mu_i}{n_i^2}}\cos(\theta_i)\sin\left(\frac{2\pi}{\lambda}n_i h_i \cos\theta_i\right) & \cos\left(\frac{2\pi}{\lambda}n_i h_i \cos\theta_i\right) \end{bmatrix} \tag{6}$$

The field strength in an n-layer system and their Fresnel behavior is

$$Q_i = M_i^{-1}\prod_{i=1}^{n-1} M_i Q_{n-1} \tag{7}$$

where is defined as the matrix of the field strength at each layer is

$$Q_i = \begin{bmatrix} H_Y^0 \\ E_Z^0 \end{bmatrix} \tag{8}$$

$$M_{ij} = \left(\prod_{k=2}^{n-1} M_k \right)_{ij} \tag{9}$$

Final equation for reflection coefficient for P-wave™ is

$$r^p = \frac{(M_{11} + M_{12} q_N) q_1 - (M_{21} + M_{22} q_N)}{(M_{11} + M_{12} q_N) q_1 + (M_{21} + M_{22} q_N)} \tag{10}$$

3 Results and Discussions

Figure 1 shows reflectance profile with thickness of gold layer for different values of incidence angle. With increase of angle, magnitude of reflectance increases. The rate of increment is higher toward the thickness of gold layer in low magnitude, as it depicts the effectiveness of the reflecting radiant energy. As reflectivity indicates the limiting value of reflectance as the sample becomes thick, in this case also after certain limit of increasing the thickness of gold layer the reflectance gets saturated.

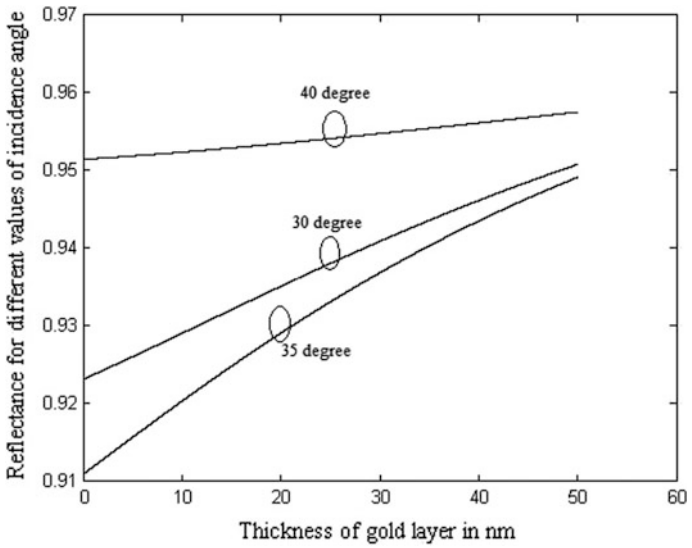


Fig. 1 Reflectance profile with gold layer thickness for different angle of incidences

Figure 2 shows the thickness of gold layer profile with reflectance for different values of wavelength of He–Ne laser light. With increase of wavelength of He–Ne laser light, magnitude of reflectance decreases while increasing the thickness of gold layer reflectance increases. For higher values of wavelength, light travels higher depth into the material so we get less radiant energy from the surface after reflection hence reducing the reflectance.

Figure 3 shows the incident angle profile with reflectance for different values of thickness of gold layer. At the angle where there is a sharp dip of reflectance is when the propagation constant of the incident wave along the surface is close to propagation constant of the surface plasmon between gold and glass medium. After certain angle of incidence the reflectance's are converging independent of the thickness of the gold layer. As also we can observe here before critical angle reflectance also decreases in a small amount. At resonance when resistance of medium is maximum then the power of the surface plasmon is lost by internal absorption in the metal, so radiative energy from the surface becoming less thus reducing the reflectance.

Figure 4 shows the incident angle profile with reflectance for different values of wavelength of He–Ne laser light. At 633 nm wavelength reflectance is zero at certain incident angle; at this wavelength frequency of the light matches with the natural oscillation frequency of the electronic charges on the metal boundary which giving rise to the fluctuations of charges thus increasing the power loss hence reducing the reflectance. Now more is the wavelength lesser is the frequency which agitates the charges in small amount thus increasing the reflectance with respect to the 633 nm wavelength.

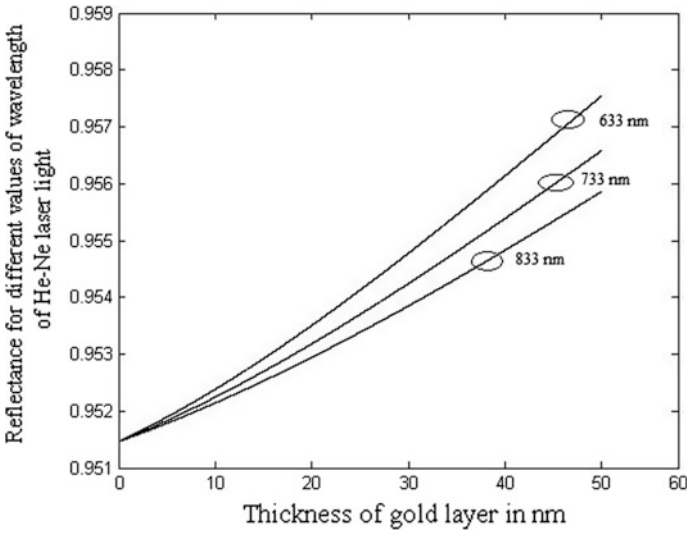


Fig. 2 Reflectance profile with gold layer thickness for different source wavelengths

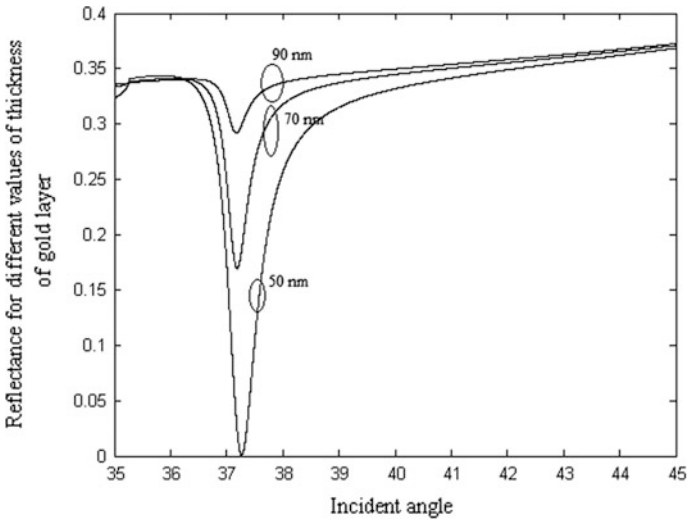


Fig. 3 Reflectance profile with incident angle for different gold layer thickness

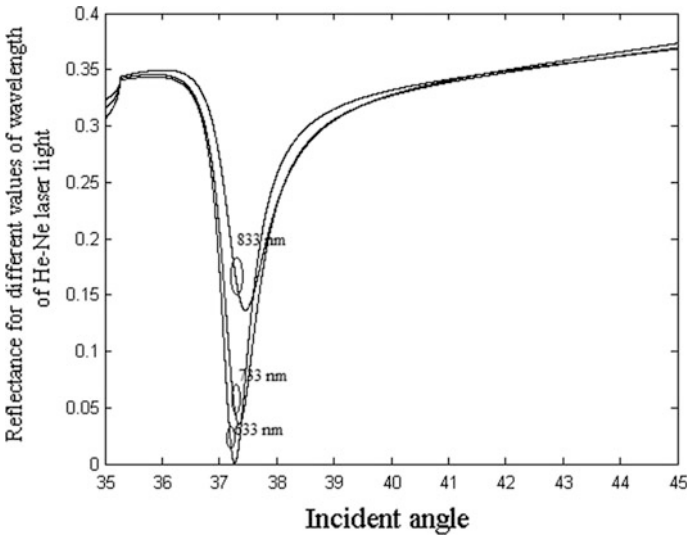


Fig. 4 Reflectance profile with incident angle for different source wavelengths

Figure 5 shows the field enhancement for three-layer profile with incidence angle. As when the incidence angle increases the field intensity at the metal surface goes on increasing. At critical incidence angle, the field intensity is at maximum at the metal surface as most of the energy passes along the surface. If again the incidence angle increases then energy starts going inside the metal, thus reducing

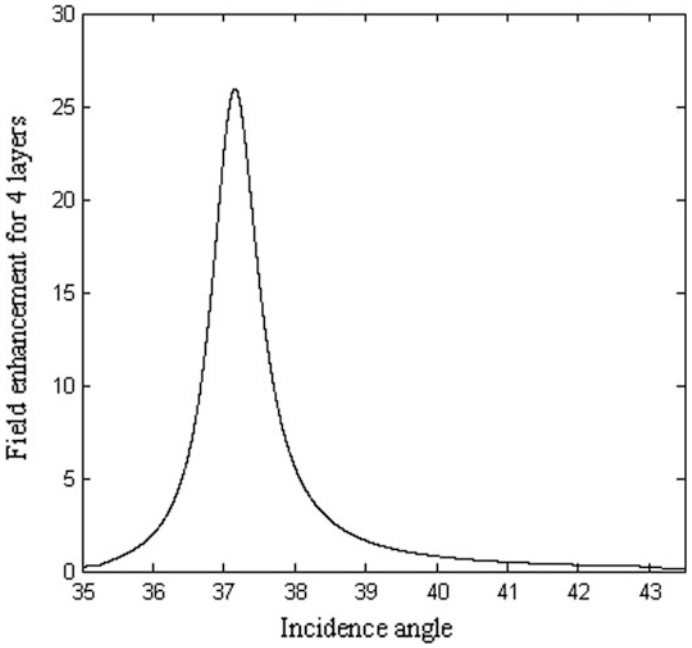


Fig. 5 Field enhancement profile with incident angle

the reflectivity and it then increases field intensity inside the metal thus reducing the field intensity at the surface. So there is the peak electric field at the surface at critical angle in the graph.

4 Conclusion

In this paper, we present the variation of reflectance for design of surface plasmon-based sensor at visible spectra. Only three-layer structure is considered, and dependence of structural and input parameters is taken into account within fabrication range to observe the variation. Generalized mathematical formulation is presented for which work can be extended into n-layer structure. Results demonstrate that critical angle is between 37° and 37.5° where resonance occurs, and corresponding maximum field enhancement occurs. Highest magnitude of reflectance is critically dependent on number of layers, and can be increased for higher layered structure. This paper provides good theoretical findings for surface plasmon-based optical sensor for bio applications.

References

1. Sharma, A.K., Jha, R., Gupta, B.D.: Fiber-optic sensors based on surface plasmon resonance: a comprehensive review. IEEE Sens. J. **7**, 1118–1129 (2007)
2. Homola, J., Yee, S.S., Gauglitz, G.: Surface plasmon resonance sensors: review. Sens. Actuators B **54**, 3–15 (1999)
3. Bhatta, D., Stadden, E., Hashem, E., Sparrow, I.J.G., Emmerson, G.D.: Multi-purpose optical biosensors for real-time detection of bacteria, viruses and toxins. Sens. Actuator. B **149**, 233–238 (2010)
4. Taylor, J.D., Linman, M.J., Wilkop, T., Cheng, Q.: Regenerable tethered bilayer lipid membrane arrays for multiplexed label-free analysis of lipid-protein interactions on poly (dimethylsiloxane) microchips using SPR imaging. Anal. Chem. **81**, 1146–1153 (2009)
5. Boozer, C., Kim, G., Cong, S., Guan, H., Londergan, T.: Looking towards label-free biomolecular interaction analysis in a high-throughput format: a review of new surface plasmon resonance technologies. Curr. Opin. Biotechnol. **17**, 400–405 (2006)
6. Ding, S.Y., Yi, J., Li, J.F., Ren, B., Wu, D.Y., Panneerselvam, R., Tian, Z.Q.: Nanostructure-based plasmon-enhanced raman spectroscopy for surface analysis of materials. Nat. Rev. Mater. **1**, 16021 (2016)
7. Li, H., Peng, W., Wang, Y., Hu, L.: Theoretical analyses of localized surface plasmon resonance spectrum with nanoparticles imprinted polymers. In: IEEE, Communications and Photonics Conference and Exhibition, pp. 1–8 (2011)
8. Ishimaru, A., Jaruwatanadilok, S., Kuga, Y.: Generalized surface plasmon resonance sensors using metamaterials and negative index materials. Prog. Electromagn. Res. **51**, 139–152 (2005)
9. Jetté-Charbonneau, S., Charbonneau, R., Lahoud, N., Mattiussi, G.A., Berini, P.: Bragg gratings based on long-range surface plasmon-polariton waveguides: comparison of theory and experiment. IEEE J. Quantum Electron. **41**, 1480–1491 (2005)
10. Popov, A.K., Myslivets, S.A., George, T.F., Shalaev, V.M.: Compensating losses in positive and negative-index metamaterials through nonlinear-optical quantum switching. In: IEEE Paper on Lasers and Electro-Optics—Pacific Rim (2007)
11. Sreekanth, K.V., Alapan, Y., ElKabbash, M., Ilker, E., Hinczewski, M., Gurkan, U.A., De Luca, A., Strangi, G.: Extreme sensitivity biosensing platform based on hyperbolic metamaterials. Nat. Mater. **15**, 621–627 (2016)
12. Liu, Y., Liu, Q., Chen, S., Cheng, F., Wang, H., Peng, W.: Surface plasmon resonance biosensor based on smart phone platforms. Sci. Rep. **5**, 12864 (2016)
13. Demetriadou, A.: The impact of natural modes in plasmonic imaging. Sci. Rep. **5**, 18247 (2015)
14. Zavats, A.V., Darmanyad, S.A., Gcrard, D., Salomnn, L., de Fornel, F.: Surface plasmon polaritons on nanostructured surfaces and thin films. In: IEEE 10th International Conference on Mathematical Methods in Electromagnetic Theory, pp. 73–78 (2005)
15. Satija, J., Shukla, G. M., Mukherji, S.: Potential of dendrimeric architecture in surface plasmon resonance biosensor. In: IEEE International Conference on Systems in Medicine and Biology, pp. 86–89 (2011)
16. Pechprasarn, S., Somekh, M.: Analyzing surface plasmon microscopy with rigorous diffraction theory. In: IEEE Conference on Functional Optical Imaging (2011)

Newton–Krylov Subspace Method to Study Structure Parameter Optimization in Rib Waveguide Communication

Sucharita Bhattacharyya and Anup Kumar Thander

Abstract In this paper, we accomplish a comprehensive study of optical waveguide modeling incorporating modal index concept in refractive index profile for two specific rib waveguide structures using fourth-order finite difference method in combination with Newton–Krylov subspace algorithm. The numerical results verify the behavior of this higher order compact (HOC) approximations for stability and convergence with least computational time. Obtained results for normalized indices and modal indices are compared with other methods to verify the accuracy and efficiency of the simple scheme used. Also variations of these indices with waveguide structure parameters help to identify their optimized values for efficient wave propagation which are found to be material dependent.

Keywords Optical waveguide · Finite difference method · Newton–Krylov subspace algorithm · Higher order compact · Structure parameter optimization

1 Introduction

Optical waveguides as a fundamental component are widely used in the transmission and processing of signals, where optical beam works as carrier. Well-defined refractive index profiles and geometric shapes help to understand the propagation properties in such waveguides more appropriately [1–9]. Here rib waveguides, relatively easy to develop are considered as this structure allows the greatest versatility in device design. Now the most basic and important step of any communication system is the modeling process which plays a very significant role by evaluating the structural design performance through wave guiding and its

S. Bhattacharyya (✉) · A.K. Thander
Department of Applied Science and Humanities, Guru Nanak Institute of Technology, Kolkata 700114, India
e-mail: sucharita.gnit@gmail.com

A.K. Thander
e-mail: anup.thander@yahoo.co.in

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_22

confinement capability. Additionally, it can reduce the high cost required during the fabrication process by optimizing the technique and design parameters, best suited the initial requirements.

Modeling can be done using analytical and numerical methods. As for these type of structures no precise analytic solution is obtained, different numerical techniques have been implemented for their study. The two most rigorous numerical methods used to solve waveguide problems are finite difference method (FDM) and finite element method (FEM),where if the governing equations and their boundary conditions are provided, it is easier to implement FDM rather than FEM. Accordingly, an accurate, higher order compact (HOC) FDM [1] is developed here for our optical waveguide system.

Now over the past decades, considerable attention has been paid to the development of software tools that can accurately simulate and evaluate the performance of such communication systems. But as optical links become more popular in communication networks involving several linear and nonlinear components, number of input parameters that can be optimized increases drastically and become very time-consuming which affect the system performance even with powerful computers. So proper optimization algorithms is employed to guide the search over the large set of input parameters to find the best solution in a minimum simulation time that converge to the optimum, reducing the number of iterations and complexity.

For that purpose, our system of resulting linear finite difference equations are solved using successive over relaxation (SOR) type of iteration [10], but due to higher divergence and run time error, a serious problem arises when applied to various wave guides. So we have used steepest descent (SD) and conjugate gradient (CG) schemes [2], belonging to well-known Newton–Krylov Subspace category. Their performances are evaluated and compared in terms of CPU time for various grid sizes for two specific rib wave guide structures, where CG scheme proves to be the most effective one showing least computational time, assuring convergence while producing similar simulation results.

Moreover, when this scheme is used to determine the fundamental TE-polarized solutions of the Helmholtz wave equation in GaAs and GeSi rib wave guide structures introducing the concept of modal index and normalized index [3], it clearly shows that how the transmission properties of the guided waves are affected through structure parameter optimization.

Following the introduction, theoretical concept and numerical simulations are given in Sect. 2. In Sect. 3, the propagating field is analyzed in terms of run time comparison with different iteration methods and structure parameter optimization of the considered rib wave-guides. Different conventional approaches show close agreement with the results obtained using our scheme. A brief conclusion is given in Sect. 4.

2 Theoretical Reviews and Numerical Computation

The configuration of the optical rib waveguide structure used by us is shown in Fig. 1 with rib width w and inner rib height H. Rib outer region has a thickness of d and h is the height of the rib, λ is the free-space optical wavelength [4]. The three regions have refractive indices of n_C, n_G, and n_S at our chosen λ value. Now considering harmonic wave propagation in the z direction along a rib waveguide, and following usual procedure we obtain the Helmholtz wave equation for quasi TE modes (with E_x continuous across horizontal interfaces but discontinuous across vertical interfaces) as the eigen solutions of the equation

$$[\nabla_T^2 + k^2]E_x = \beta^2 E_x \tag{1}$$

with ∇_T^2 as the Laplacian operator acting on x and y and

$$k(x,y) = \omega(\varepsilon\mu)^{\frac{1}{2}} = \frac{2\pi n(x,y)}{\lambda} \tag{2}$$

is the total propagation constant. Here ω is the angular frequency and $n(x,y)$ is refractive index, $\beta^2$is the eigen value corresponding to the eigen function E_x. Here dielectric constant $\varepsilon(x,y)$ is piecewise constant throughout the solution domain, permeability μ is completely constant and x and y represent the coordinates of the transverse section of the waveguide.

In order to solve Eq. (1) numerically by compact fourth-order finite difference method, we have decomposed the rectangular solution domain into uniform grid points whose x length and y length are h_x and h_y respectively as shown in Fig. 2.

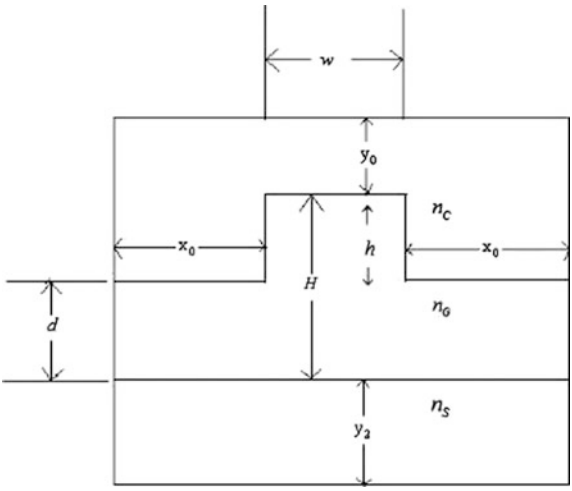


Fig. 1 Rib waveguide structure

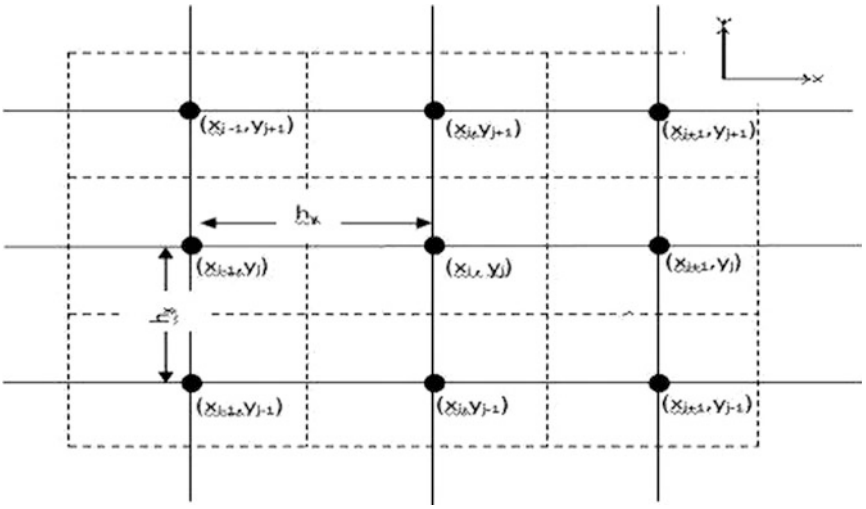


Fig. 2 Nine point stencil of the field domain

Then following the scheme developed in [1, 4], a simultaneous system of linear algebraic equation with nine-diagonal coefficient matrix is obtained as

$$\sum_{\substack{k=-1,0,1\\ l=-1,0,1}} A_{i,j}^{k,l} E_{i+k,j+l} = 0 \tag{3}$$

where

$$A_{i,j}^{1,1} = a_{i,j}\left[\frac{2k_{i+1,j+1}^2}{k_{i,j+1}^2 + k_{i+1,j+1}^2}\right], \quad A_{i,j}^{1,-1} = b_{i,j}\left[\frac{2k_{i+1,j-1}^2}{k_{i+1,j}^2 + k_{i+1,j-1}^2}\right],$$

$$A_{i,j}^{-1,1} = c_{i,j}\left[\frac{2k_{i-1,j+1}^2}{k_{i-1,j}^2 + k_{i-1,j+1}^2}\right], \quad A_{i,j}^{-1,-1} = d_{i,j}\left[\frac{2k_{i-1,j-1}^2}{k_{i,j-1}^2 + k_{i-1,j-1}^2}\right],$$

$$A_{i,j}^{1,0} = b_{i,j}\left[\frac{k_{i+1,j-1}^2 - k_{i+1,j}^2}{k_{i+1,j-1}^2 + k_{i+1,j}^2}\right] + q_{i,j}\left[\frac{2k_{i+1,j}^2}{k_{i,j}^2 + k_{i+1,j}^2}\right],$$

$$A_{i,j}^{-1,0} = c_{i,j}\left[\frac{k_{i-1,j+1}^2 - k_{i-1,j}^2}{k_{i-1,j+1}^2 + k_{i-1,j}^2}\right] + f_{i,j}\left[\frac{2k_{i-1,j}^2}{k_{i,j}^2 + k_{i-1,j}^2}\right],$$

$$A_{i,j}^{0,1} = a_{i,j}\left[\frac{k_{i+1,j+1}^2 - k_{i,j+1}^2}{k_{i+1,j+1}^2 + k_{i,j+1}^2}\right] + g_{i,j}\left[\frac{2k_{i,j+1}^2}{k_{i,j}^2 + k_{i,j+1}^2}\right],$$

$$A_{i,j}^{0,-1} = d_{i,j}\left[\frac{k_{i-1,j-1}^2 - k_{i,j-1}^2}{k_{i-1,j-1}^2 + k_{i,j-1}^2}\right] + h_{i,j}\left[\frac{2k_{i,j-1}^2}{k_{i,j}^2 + k_{i,j-1}^2}\right],$$

$$A_{i,j}^{0,0} = q_{i,j}\left[\frac{k_{i+1,j}^2 - k_{i,j}^2}{k_{i+1,j}^2 + k_{i,j}^2}\right] + f_{i,j}\left[\frac{k_{i-1,j}^2 - k_{i,j}^2}{k_{i-1,j}^2 + k_{i,j}^2}\right] + g_{i,j}\left[\frac{k_{i,j+1}^2 - k_{i,j}^2}{k_{i,j+1}^2 + k_{i,j}^2}\right] + h_{i,j}\left[\frac{k_{i,j-1}^2 - k_{i,j}^2}{k_{i,j-1}^2 + k_{i,j}^2}\right] - m_{i,j}$$

$$a_{i,j} = 1 - \left[\frac{h_x^2 h_y^2}{12(h_x^2 + h_y^2)}\right](\beta^2 - k_{i,j}^2), b_{i,j} = 1 - \left[\frac{h_x^2 h_y^2}{12(h_x^2 + h_y^2)}\right](\beta^2 - k_{i,j}^2),$$

$$c_{i,j} = 1 - \left[\frac{h_x^2 h_y^2}{12(h_x^2 + h_y^2)}\right](\beta^2 - k_{i,j}^2), d_{i,j} = 1 - \left[\frac{h_x^2 h_y^2}{12(h_x^2 + h_y^2)}\right](\beta^2 - k_{i,j}^2),$$

$$q_{i,j} = \frac{12h_y^2}{h_x^2 + h_y^2} - 2 - \frac{5h_x^2 h_y^2}{6(h_x^2 + h_y^2)}(\beta^2 - k_{i,j}^2), q_{i,j} = \frac{12h_y^2}{h_x^2 + h_y^2} - 2 - \frac{5h_x^2 h_y^2}{6(h_x^2 + h_y^2)}(\beta^2 - k_{i,j}^2),$$

$$g_{i,j} = \frac{12h_x^2}{h_x^2 + h_y^2} - 2 - \frac{5h_x^2 h_y^2}{6(h_x^2 + h_y^2)}(\beta^2 - k_{i,j}^2), h_{i,j} = \frac{12h_x^2}{h_x^2 + h_y^2} - 2 - \frac{5h_x^2 h_y^2}{6(h_x^2 + h_y^2)}(\beta^2 - k_{i,j}^2),$$

$$m_{i,j} = 20 - \frac{77h_x^2 h_y^2}{3(h_x^2 + h_y^2)}(\beta^2 - k_{i,j}^2)$$

which can be rewritten in its usual form of

$$A_{TE}E_{TE} = \beta_{TE}^2 E_{TE} \tag{4}$$

where A_{TE} is a real nonsymmetric band matrix with β_{TE}^2 is TE propagation eigen value and E_{TE} is the corresponding normalized eigenvector representing the field profile E_x. Here, we have used zero field outer boundary condition.

A. Rayleigh Quotient Solution (RQS)

To find out any modal eigenvalue β^2 contained in Eq. (4), we use classical Rayleigh Quotient Solution [6] of the matrix eigenvalue problem as

$\beta^2 = \frac{<E_{TE}^T A_{TE} E_{TE}>}{<E_{TE}^T E_{TE}>}$ where the superscript T denotes the transpose of a matrix. After an estimation of β^2, the field profile E_{TE} is updated from an iterative solution of the system of linear algebraic equations (4) by CG Method, as it has least computational time compared to SOR and SD method and our coefficient matrix system is nonsymmetric with a possibility of having complex eigen values [2].

B. Normalized Index

It is very important to mention that we have applied here the effective index concept for the guiding region of the rib waveguide as there are discontinuities existing in the dielectric interfaces and there must be an average value for the refractive index which would reduce a layered waveguide to an equivalent uniform one [3]. This effective index value can be used to calculate normalized index $b = \frac{n_{eff}^2 - n_s^2}{n_G^2 - n_s^2}$ which indicates that how far a mode is from cut off, playing a very important role so far as propagation is concerned through a waveguide. For propagation to be possible, it is required that $0 \leq b < 1$.

3 Results and Discussion

The scheme developed in [1] along with RQS technique, is used here successfully to calculate effective refractive index and normalized index for two different dielectric rib waveguides mostly used in optoelectronic integrated circuits, the GeSi-Si heterostructure and GaAs/GaAlAs system with air cladding [5–7]. The geometrical and optical parameters of the structures S2 and S4 [7] in GaAs and GeSi are shown in Table 1.

Here, it may be noted that in order to achieve high coupling efficiency with the standard single-mode fiber, the total rib height (H) of the rib waveguide is set to be 1.0 μm, and the operating wavelength is selected at 1.3 μm for S2 structure and at 1.15 μm for S4 structure. Also, Rib outer region thickness d is considered to vary from 0.6 to 1 μm with the rib height varying correspondingly from 0.4 to 0 μm.

As already mentioned above we have studied all the three iterative methods for our system of equations to determine the electric field profile. The corresponding computational (CPU) time (in sec) for various grid sizes are calculated and shown in comparison Table 2 which are obtained using MATLAB on a computer having processor configuration: Intel Core i5 3550 @ 3.3 Ghz.

From the table, it is clear that the CG Method takes least computational time with respect to the other two and shows regular convergence as expected [2]. So, here we have used CG algorithm as the most appropriate iterative scheme to solve (4) and the corresponding code in MATLAB has been developed for our nine-diagonal coefficient matrix system to determine the electric field profile, effective index n_{eff}, and the normalized index b through structure parameter optimization.

The calculated values for the rib waveguide structures under consideration are shown in Table 3, where GaAs results are compared with other available data including those of Stern. Comparisons show very strong agreement (within 1%) for both the indices.

Additionally, We have studied the variation of n_{eff} and b with outer rib thickness d within the ranges from 0.6 to 1 μm with corresponding changes in rib height h from 0.4 to 0 μm for GaAs rib waveguide and is shown in Table 4. It is clear from the table, that as d increases, n_{eff} and b increases. Different workers also got the similar results as shown below. Actually, higher d value should enhance the

Table 1 Optical and geometrical structure parameters (in μm) of GaAs and GeSi rib waveguides [5, 7]

Structure	n_G		n_S		n_C		H	h	d	w	λ
	GaAs	GeSi	GaAs	GeSi	GaAs	GeSi					
S2	3.44	3.6	3.36	3.5	1.0	1.0	1.0	0.1	0.9	3	1.3
S4	3.44	3.6	3.40	3.5	1.0	1.0	1.0	0.4–0.0	0.6–1.0	3	1.15

Table 2 CPU time comparison for various iterative methods in S2 and S4 structures

h_x	h_y	S2 Structure			S4 Structure		
		SOR	SDM	CGM	SOR	SDM	CGM
0.058	0.09	5079.29895	331.595725	8.86085679	5292.30272	831.984533	3.40082179
0.06	0.09	4514.87174	250.678006	13.4784863	733.984704	99.7158391	3.32282130
0.07	0.09	3641.09454	265.966104	47.22150270	563.569212	38.6414476	9.84366310
0.08	0.09	2779.89101	195.282051	3.96242540	553.943950	4.39922819	8.28365309
0.08	0.092	2804.47677	291.721870	4.83603099	671.631105	593.623887	12.5580804
0.078	0.09	3130.26926	3.74402399	1.65361059	612.288324	565.61282	13.7124878

Table 3 Comparison of modal and normalized index for S2 and S4 structure of GaAs

Structures	S2		S4 (for d = 0.9)	
Method	n_{eff}	b	n_{eff}	b
Present Method [1]	3.3958327	0.4450	3.4159040	0.3962
Stern [7]	3.3953942	0.4395	3.41568	0.3905
FEM [8]	3.3948874	0.4332	3.4155996	0.3886
BPM [8]	3.3944707	0.4280	3.4151870	0.3783
VPM [8]	3.3950156	0.4348	3.4153312	0.3819

Table 4 Variation of normalized and modal index with d and comparison with other available data for S4 structure of GaAs waveguide

d	Vfem [8]		Fem [8]		Stern [7]		Present Method [1]	
	n_{eff}	b	n_{eff}	b	n_{eff}	b	n_{eff}	b
0.6	3.41357244	0.338	3.413648582	0.3399	3.41353	0.3368	3.413616522	0.3391
0.7	3.414053309	0.35	3.414101393	0.3512	3.41406	0.3503	3.414009233	0.3489
0.8	3.414734426	0.367	3.414838585	0.3696	3.41472	0.3665	3.414874639	0.3705
0.9	3.415575618	0.388	3.415599649	0.3886	3.41554	0.3872	3.415904027	0.3962
1	3.417137281	0.427	3.417101251	0.4261	3.41716	0.4275	3.417261383	0.4301

confinement of propagating e.m. wave in the guiding zone of the waveguide which is confirmed from the corresponding b values indicating the presence of mode far away from cut off. This means vertical spread decreases signifying better confinement of the propagating wave.

Now in order to obtain optimization of wave guide section dimensions for GaAs and GeSi rib wave guide, a comparative study is shown in Table 5 for the variation of modal index and normalized index for S2 structure of GaAs and GeSi rib wave guide with d. Here, different b values obtained for same d and λ (1.3 μm) in GaAs and GeSi wave guide clearly shows material dependence of propagation property which holds equally well for other wavelengths too. Here, it is noted that for $d \leq 0.8$ μm, b values for GaAs is smaller than that for GeSi in each case, but the situation reversed for $d \geq 0.9$ where smaller vertical spread (as b is larger) is obtained for GaAs than GeSi showing better confinement capability for the former one within the specified range.

Similarly, Table 6 presents variation of b and n_{eff} with w, the rib width ranging from 2.55 to 4.0 μm where it is found that b values for GaAs is smaller than for GeSi for $w \leq 2.85$ μm but for $w \geq 3$, normalized index value b increases for GaAs compared to GeSi showing better confinement capability for GaAs in this range. Here, it may be noted that too small width (w) will reduce the restriction of the rib structure on the guiding mode, whereas too big width (w) will be difficult to tackle for system integration (as then the wave guide bends and branches), the width (w) of our rib wave guide is set to be in the range of 2–5 μm.

Table 5 Comparison of modal and normalized index values for S2 structure in GaAs and GeSi material systems varying d

	GaAs		GeSi	
d	n_{eff}	*b*	n_{eff}	*b*
0.6	3.3644678441	0.0552	3.511346362	0.1120
0.7	3.3811742288	0.2624	3.529157718	0.2887
0.8	3.3909078994	0.3836	3.539549032	0.3921
0.9	3.3958346576	0.4450	3.544838627	0.4449
1.0	3.3985031848	0.4784	3.547798297	0.4745

Table 6 Comparison of modal and normalized index values for S2 structure in GaAs and GeSi material system varying w

	GaAs		GeSi	
w	n_{eff}	*b*	n_{eff}	*b*
2.55	3.394928387593926	0.4338	3.543825102	0.4348
2.70	3.394943401217174	0.4339	3.543841395	0.4349
2.85	3.395331121614340	0.4388	3.544275005	0.4392
3.00	3.395834657602020	0.4450	3.544838627	0.4449
3.50	3.396337940404861	0.4513	3.545401045	0.4505
4.00	3.397016869103838	0.4598	3.546160826	0.4581

Now the interesting point to mention here is that the optimized values chosen by Stern [5] for d is 0.9 μm and for w as 3 μm in case of S2 structure with $\lambda = 1.30$ μm for GaAs. So the optimized result obtained for normalized index b within the specified ranges of d and w for GaAs using our scheme completely agree with the prediction of Stern. So based on the analysis of the normalized index value, the cross-sectional dimensions of the GaAs as well as GeSi rib waveguides are optimized here through FDM simulation and CG iteration technique though the results for GeSi could not be verified due to unavailability of data.

4 Conclusions

In this work, the effective refractive index and normalized index of two different optical rib waveguide structures are obtained by higher order compact finite difference method using conjugate gradient iteration scheme for GaAs and GeSi material systems and the results are found to be in close agreement with other published results. Most importantly, it is to be noted that our method successfully verifies the significant role of structure parameter optimization in propagating e.m. wave which could be taken into account during the fabrication process of optical waveguides most commonly used in miniaturization and integration of the communication system.

Acknowledgements The authors are very grateful to University Grants Commission (Grant No. F.PSW-180/13-14(ERO)), Govt. of India for providing the necessary research fund to carry out this research work and to the JIS Group Educational Initiatives to provide the necessary infrastructure.

References

1. Thander, A.K., Bhattacharyya, S.: Study of optical wave guide using HOC scheme. Appl. Math. Sci. **8**(79), 3931–3938 (2014)
2. Bhattacharyya, S., Thander, A.K.: Optical wave guide analysis using higher order compact FDM in combination with Newton Krylov Subspace methods. In: 2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks, pp. 66–71 (2015)
3. Thander, A.K., Bhattacharyya, S.: Rib wave guide propagation through modal index study using higher order compact (HOC) finite difference method (FDM). In: 2015 6th International Conferences on Computers and Devices for Communications (CODEC-2015) (2015)
4. Thander, A.K., Bhattacharyya, S.: Optical confinement study of different semi conductor rib wave guides using higher order compact finite difference method. Optik **127**, 2116–2120 (2016)
5. Stern, M.S.: Semivectorial polarised finite difference method for optical waveguides with arbitrary index profiles. IEE Proceedings-Optoelectron **135**(1), 56–63 (1988)
6. Soref, R.A., Schmidtchen, J., Petermann, K.: Large single-mode rib waveguides in GeSi-Si and Si-on-Sio2. IEEE J. Quantum Electron. **27**(8) (1991)
7. Robertson, M.J., Ritchie, S., Dayan, P.: Semiconductor waveguides: analysis of optical propagation in single rib structures and directional couplers. IEE Proc. J. **132**, 336–342 (1985)
8. Huang, Weiping, Hauss, Hermann A.: A simple variational approach to optical rib waveguides. J. Lightwave Technol. **9**(1), 56–61 (1991)
9. Rahman, B.M.A.: Finite element analysis of optical waveguides. In: Prog. Electromagnet. Res. PIER **10**, 187–216 (1995)
10. Hoffman, J.D.: Numerical Methods for Engineers and Scientists. Marcel Dekker, Inc., New York (1992)

Fuzzy-Tuned SIMC Controller for Level Control Loop

Ujjwal Manikya Nath, Chanchal Dey and Rajani K. Mudi

Abstract Internal model control (IMC) technique is one of the well-accepted model-based controller designing methodologies which is widely used in process industries due to their simplicity and ease of tuning. Most of the IMC tuning provides good set point response but unsatisfactory load rejection behavior. To overcome this limitation for industrial processes SIMC technique is reported in the literature. In this technique, to derive the SIMC controller expression, higher order processes are approximated as first-order plus time delay model. Hence, uncertainty is always there in process modeling and as a result SIMC controller may fail to provide the satisfactory performance with conventional fixed tuning. A fuzzy-tuned SIMC controller is reported here to surmount this drawback and its efficacy is established through real-life experimentation on a laboratory-based level control loop.

Keywords SIMC controller · Fuzzy auto-tuner · Model identification · Level control process

U.M. Nath (✉) · R.K. Mudi
Department of Instrumentation and Electronics Engineering,
Jadavpur University, Kolkata 700098, India
e-mail: um.nath@yahoo.com

R.K. Mudi
e-mail: rkmudi@yahoo.com

C. Dey
Department of Applied Physics Instrumentation and Control Engineering,
University of Calcutta, 92 A.P.C. Road, Kolkata 700009, India
e-mail: chanchaldey@yahoo.co.in

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_23

1 Introduction

Model-based tuning methodology has now become a well-accepted practice for industrial PID controller tuning [1]. Among the available model-based controller designing techniques internal model control (IMC) [2] method is very popular due to its simplicity and straightforward approach. The most important feature of IMC tuning is that it has a single tuning parameter. IMC-based PID controller (IMC-PID) is primarily reported by Rivera et al. [3] for first-order single-input single-output (SISO) processes, and later it is also extended for multivariable processes [4, 5]. By choosing an appropriate value of the sole tuning parameter, i.e., closed-loop time constant (λ) of IMC controller desired process response can be obtained. To achieve improve load rejection behavior Skogestad [6] proposed a modified IMC (SIMC) method for processes with large time constant. Usually, IMC provides good set point tracking for larger value of λ, whereas smaller value of λ is essential to obtain superior disturbance rejection. Hence, instead of having a fixed tuning for λ, an auto-tuning feature is incorporated here with SIMC structure so that an overall improved process responses can be achieved. Nine fuzzy rules are designed for online tuning of λ depending on the current process error (e) and change of error (Δe). An online adaptive IMC controller is first reported in [7] and later [8, 9] described IMC adaptation schemes for specific class of processes.

Here our main objective is to design an auto-tuning scheme for a well-accepted SIMC controller setting [6] so that an improved process behavior can be obtained during both the transient as well as steady state responses. Real-time experimental verification of the proposed scheme is made on a tank level process by Feedback Instrument Ltd. [10] under set -point change and load variation. Process responses along with quantitative performance indices (IAE and ITAE) clearly justify the efficacy of the reported controller over the conventional fixed tuning (SIMC) [6].

2 Proposed Auto-Tuned Controller Design

2.1 SIMC Technique

In [6] Skogestad proposed a very simple IMC technique termed as SIMC which is very easy to implement for closed -oop applications. In SIMC, the suggested value of closed-loop time constant is considered to be the dead time of the concerned process, i.e., $\lambda = \theta$. But, it is found that though this setting provides good transient response during set point tracking but offers sluggish recovery after load change. To overcome this limitation Skogestad suggested for a modification of the integral time for processes with large time constant. He proposed that the integral time to be considered as $\tau_I = \min\{(\tau_p + 0.5\theta), 4(\theta + \lambda)\}$ to get improved load regulation. Hence, considering the prior guideline for closed-loop time constant $\lambda = \theta$ the expression for integral time can be written as $\tau_I = \min\{(\tau_p + 0.5\theta), 8\theta\}$.

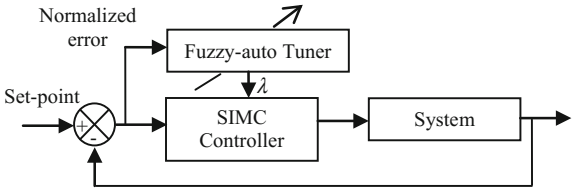


Fig. 1 Block diagram of fuzzy-tuned IMC-PI controller

2.2 Proposed Fuzzy Rule-Based Auto-Tuned SIMC Controller

In our proposed method we want to provide an auto-tuning facility for providing continuous variation of λ based on a fuzzy-rule base consisting of nine rules only. In the proposed fuzzy tuner appropriate value of λ is provided by fuzzy rules based on the current process operating conditions, i.e., process error (e) and change of error (Δe). Here, we used Mamdani [11] inference technique with gravity method of defuzzification. Block diagram of the proposed auto-tuned SIMC controller is shown in Fig. 1.

2.3 Membership Function

In our proposed method, we considered three triangular membership functions (MFs) [12] for the input variables e and Δe defined as Negative (N), Zero (Z), and Positive (P) as shown in Fig. 2a. They are defined over the normalized domain -1 to 1. Similarly, three triangular membership functions (MFs) are also defined for the output variable λ as Small (S), Medium (M), and Large (L) over the range 0–2θ (where θ is the process dead time) as defined in Fig. 2b. All the MFs are symmetrical triangle in shape and 50% overlapped with neighboring MFs.

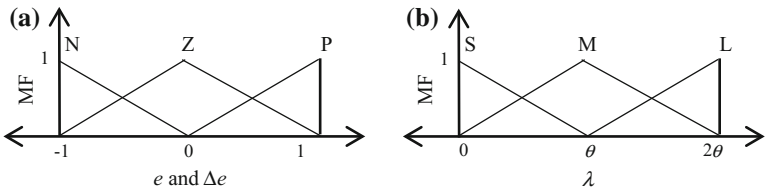


Fig. 2 Triangular MFs for **a** inputs e and Δe with 3 fuzzy **b** corresponding λ

Δe \ e	N	Z	P
N	L	L	S
Z	L	M	S
P	L	S	S

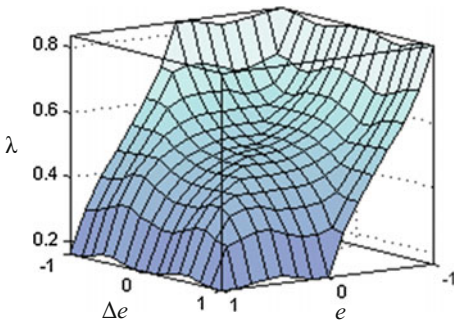


Fig. 3 Rule bases and control surfaces for 9 rules

2.4 Rule Base

The rule for providing λ is given by [13]

'IF e is E AND Δe is ΔE THEN λ is $\Delta\lambda$' (Fig. 3).

3 Model Identification

Our experimental setup is a laboratory scale level control loop manufactured by Feedback Instruments Ltd [10]. To design the IMC controller, we have to first identify the process model. Open-loop step test is performed to obtain the process reaction curve (PRC) [14] as shown in Fig. 4. From the PRC all process parameters are calculated and identified first order with time delay is model is given by the following relation: $G(s) = \frac{8.4}{42.03s+1} e^{-3s}$.

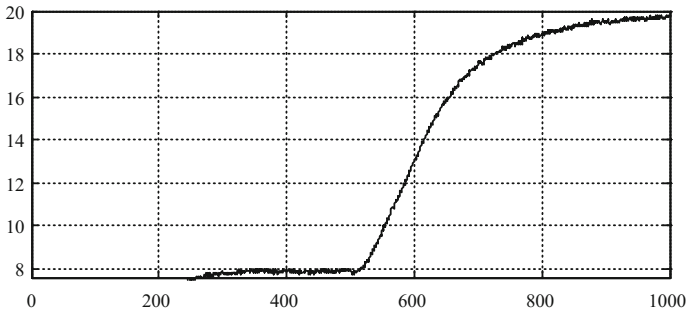


Fig. 4 PRC response of the tank

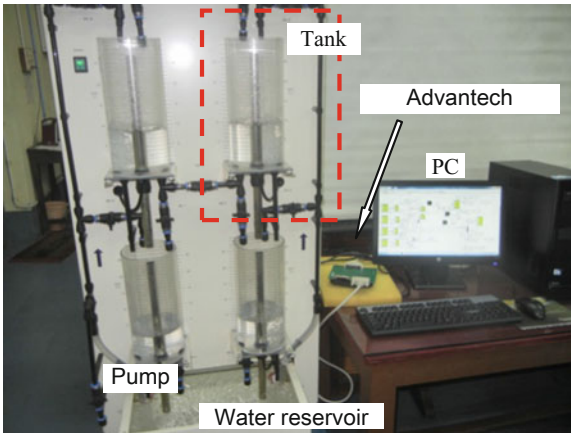


Fig. 5 Level control of tank

4 Real-Time Experimentation

Experimental setup is shown in Fig. 5 which is a quad-coupled tank process manufactured by Feedback Instrument Ltd. [10]. For experimentation we choose the right-hand side upper tank (boundary line of Fig. 5) and here our goal is to maintain the liquid level at the desired value irrespective of set point change and load disturbances. Interface between the PC and experimental setup is performed by Advantech PCI 1711 data acquisition interface card. Our controller is designed in PC in MATLAB/SIMULINK environment. Final control element (FCE) is the variable drive of a submerge pump. So the control action (0–5 V) goes to that pump through drive circuit so that the pumping rate can be regulated. Tank level signal is observed during set point tracking and under two separate load disturbances occurring in opposite directions as shown in Fig. 6.

5 Results

Responses of fuzzy-tuned SIMC controllers with nine rules are shown in Fig. 6 and the related performance indices are depicted in Table 1. The responses for SIMC and fuzzy-tuned SIMC with nine rule base are plotted in dotted black line and solid red line, respectively. Integral absolute error (IAE) and integral time absolute error (ITAE) [14] are calculated for quantitative performance analysis. From Table 1 it is found that the values of IAE and ITAE are lesser in case of our proposed fuzzy tuner compared to conventional SIMC method.

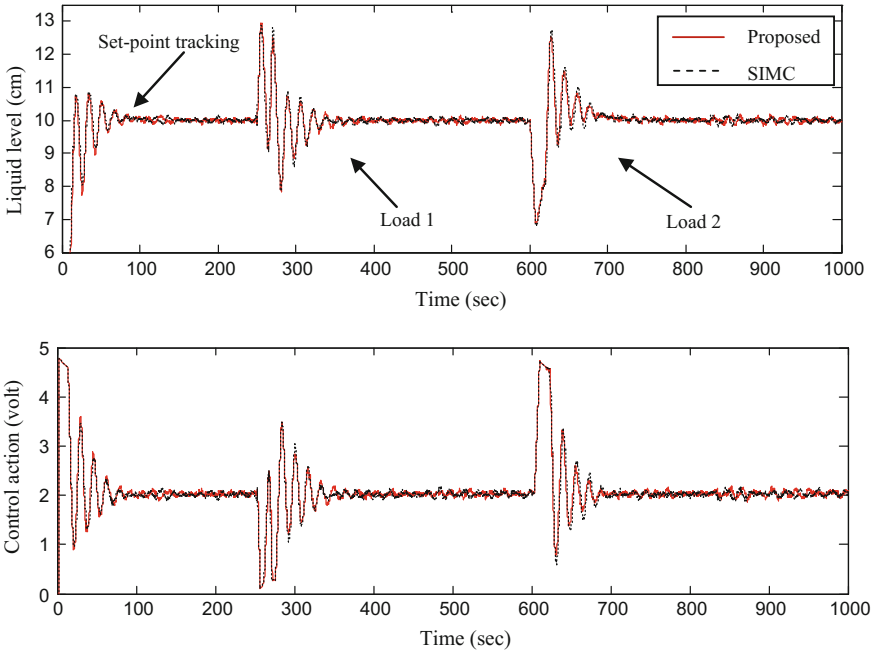


Fig. 6 Performance analysis of the SIMC and proposed fuzzy-tuned SIMC controllers under set point change and load variation

Table 1 Performance indices

Controller mode	IAE			ITAE		
	Total	Load 1	Load 2	Total	Load 1	Load 2
SIMC	336.40	82.91	107.01	9.85×10^4	2.53×10^4	7.04×10^4
Auto-tuned SIMC	332.70	79.81	102.80	9.48×10^4	2.40×10^4	6.79×10^4

6 Conclusion

SIMC controller is widely accepted in industry due to its simple designing structure and easy to implement. But, with a fixed value of the tuning parameter it fails to provide improved process responses under set point change and load variation simultaneously. To overcome this limitation a simple fuzzy-based tuner is proposed here which provides the appropriate value of the closed-loop time constant depending on the current process operating conditions. Hence, an overall improvement in process response is obtained as depicted by the experimental study. In designing the fuzzy tuner only nine rules are used, so there is a scope for further fine tuning of the SIMC controller with larger number of rules so that further enhancement in process responses can be realized.

References

1. Garcia, G.E., Morari, M.: Internal model control—1 A unifying review and some new results. Ind. Eng. Chem. Process Des. Dev. **21**, 308–323 (1982)
2. Nath, U.M., Datta, S., Dey, C.: Centralized auto-tuned IMC-PI controllers for a real time coupled tank process. Int. J. Sci. Technol. Manage. **4**(1), 1094–1102 (2015)
3. Rivera, D.E., Skogested, S., Morari, M.: Internal model control for PID controller design. Ind. Eng. Chem. Process Des. Dev. **25**, 252–265 (1986)
4. Nath, U.M., Datta, S., Dey, C.: Centralized auto-tuned IMC-PI controllers for industrial coupled tank process with stability analysis. In: 2nd International Conference on Recent Trends in Information Systems. IEEE (2015)
5. Datta, S., Nath, U.M., Dey, C.: Design and Implementation of Decentralized IMC-PI Controllers for Real Time Coupled Tank Process. MFIIS, Michael Faraday IET International Summit (2015)
6. Skogestad, S.: Analytic rules for model reduction and PID controller tuning. J. Process Control **13**, 291–309 (2003)
7. Datta, A., Ochoa, J.: Adaptive internal model control: design stability and analysis. Automtica **32**, 261–266 (1996)
8. Rupp, D., Guzzella, L.: Adaptive internal model control with application to fueling control. Control Eng. Pract. **18**(8), 873–881 (2010)
9. Silva, G.J., Datta, A.: Adaptive internal model control: The discrete-time case. Int. J. Adapt. Control Signal Process. **15**(1), 15–36 (2001)
10. Feedback Instruments Ltd. East Sussex, UK
11. Lee, C.C.: Fuzzy logic in control systems: fuzzy logic controller, part II. IEEE Trans. Syst. Man Cybern. **20**(2), 419–435 (1990)
12. Dey, C., Mudi, R.K., Mitra, P.: A self-tuning fuzzy PID controller with real-time implementation on a position control system. In: Emerging Applications of Information Technology (EAIT). IEEE (2012)
13. Nath, U.M., Dey, C, Mudi, R.K.: Fuzzy-based Adaptive IMC-PI controller for Real-time application on level control loop. In: Springer-FICTA (2016)
14. Seborg, D.E., Edgar, T.F., Melichamp, D.A.: Process Dynamic and Control, 2nd edn. Wiley (2004)

Utilization of Electromagnetic Sensor for Structural Characterization of Steels During Processing and in-Service Components

Rajat K. Roy, M. Premkumar, Ashis K. Panda and Amitava Mitra

Abstract An electromagnetic sensing device, *MagStar*, has been developed by authors, measure both magnetic hysteresis loop and magnetic Barkhausen emission. The device explains its credibility for characterizing different types steels, SA213T22 and IF steel with varying microstructural phases treated at different heat treatment conditions. *MagStar* is also very versatile to be applicable for both industries and educational research due to its portability and low cost.

Keywords Steel structures · Magnetic parameters · Electromagnetic sensor

1 Introduction

The understanding of material properties at different time period has an increasing demand not only for industrial structure and components but also for the steel at different processing stage. This is essentially important for reducing maintenance cost of any structure/components and production cost of material processing. The continuous monitoring through nondestructive evaluation (NDE) protects the premature failure of industrial structures like boiler tubes and pipes of power plant, pipelines of petrochemical industries, etc. The examination of steel properties at different processing stage also improves the product quality.

Maximum industrial structures and components are made of steel which are magnetic in nature, and it is well established that the magnetic parameters of steel components have a good correlation to structure and mechanical properties. Since last decade, a numbers of research studies have proved that electromagnetic properties of steel have a good correlation with microstructural phase fraction [1], precipitates [2] along with the change of stress [3], temperature, and processing mechanism [4]. Therefore, the electromagnetic sensor is taking a prime role for

R.K. Roy (✉) · M. Premkumar · A.K. Panda · A. Mitra
NDE & Magnetic Materials Gr., MST Division, CSIR-National Metallurgical Laboratory, Jamshedpur 831007, India
e-mail: rajatroy.k@gmail.com; rajat@nmlindia.org

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_24

microstructural characterization many important industrial steels, such as P91, P22, austenitic stainless steel, which are applicable in many industrial structures [5]. Present authors have also developed an electromagnetic sensor named as '*MagStar'* for measurement of magnetic properties of steel components not only for the laboratory scale but also applicable in structural industries. In this research we have discussed the instrument for structural phase determination during in service condition and processing of steel.

2 Principle of Magnetic Sensor

The principle of electromagnetism was applied to develop an electromagnetic magnetic sensor *MagStar*, which is explained in elsewhere [6]. On the basis of electromagnetism principle, with increasing external magnetic field, a ferromagnetic material is magnetized due to nucleation, and growth of domains. Simultaneously, with decreasing of magnetic field, the material is demagnetized with the reversion magnetic domain movement. It occurs due to the nucleation and motion of magnetic domain walls. As a consequence, the magnetic hysteresis loop (MHL) is generated. The domain wall movement can be pinned by precipitates and lattice defects existing in material, which causes their discontinuous motions. As a result of this kind of domain motion, a pulsating magnetization can be obtained corresponding to the rapid change of magnetic flux, which phenomenon is called Barkhausen effect and known as magnetic Barkhausen emission (MBE) [7]. Since both MHL and MBE are the resultant of magnetic domain movement, the measured parameters of these techniques are also very sensitive to microstructural features such as precipitates and lattice defects in material as mentioned above. Therefore, the *MagStar* generated parameters, such as coercivity, permeability, magnetic flux density, remanence, MBE voltage, can be correlated to many microstructural phenomena of steel samples. It is reported MBE voltage of polycrystalline materials is strongly dependent on the size of grain precipitates [8].

3 Experimentation

The sensitivity and function of *MagStar* sensing device were evaluated through the measurement of MBE and MHL parameters in two different type steels SA213T22 and IF steel. The former one is applicable as boiler tubes in thermal power plant while the latter one is used as car body manufacturing in automobile industries. The detailed compositions of two steels are explained in Table 1. The samples used in this research were cut from SA213T22 tube and IF steel sheet with the sizes of $70 \times 30 \times 5$ and $70 \times 30 \times 1$ mm^3, respectively. The samples from SA213T22 were water quenched after solutionizing at 1050 °C for 1 h and followed by tempering at temperatures of 200, 600, 650, 700, 720, 750, and 800 °C. On the other

Table 1 Chemical compositions of SA213 and IF steel used in this research

Steel	Compositions (wt%)
SA213	C: 0.15, Mn: 0.3–0.6, Si-0.5, P: 0.025, S: 0.025, Mo: 0.87–1.13, Cr: 1.9–2.6, Fe-rest
IF steel	C: 0.003, Mn: 0.11, Cu-0.01, Al: 0.045, Ti: 0.047, N: 0.003, Fe-rest

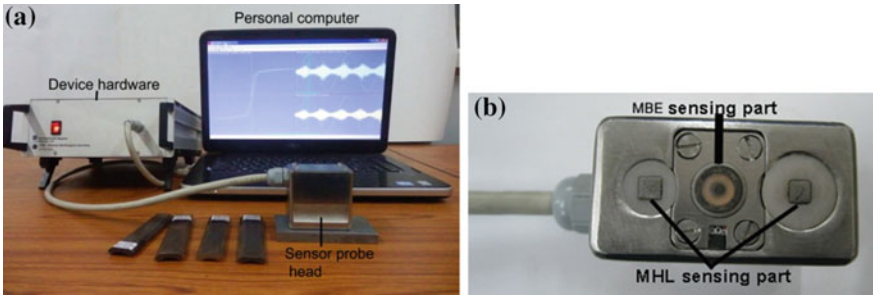


Fig. 1 Electromagnetic sensor *MagStar* **a** device hardware **b** sensor probe head

hand, the as-received IF steel was industrially cold rolled sheet with a thickness of 1 mm. The samples were annealed at the temperatures of 550, 600, 650, and 700 °C for 5 min to study the recrystallization behavior. Two different steels were heat treated at two different conditions to generate various phases for understanding the sensitivity of *MagStar* (Fig. 1) on microstructural phase evaluation. For MBE measurement, a sinusoidal cyclic magnetic field of 150–500 Oe with an excitation frequency of 40 Hz was induced, and the Barkhausen signals were filtered with a wide band pass filter (30–300 kHz). On the other hand, a sinusoidal cyclic magnetic field of 1500 Oe with an excitation frequency of 50 MHz was induced for MHL measurement.

4 Results and Discussion

4.1 Sa213 T22

Microstructural evaluation: Figure 2 shows typical optical microstructures of water quenched (WQ) and tempered SA213T22 steel. The WQ sample shows fine martensite laths (Fig. 2a). Upon progress of tempering, the martensite becomes stress free and converts to bainite and ferrite laths at the different temperature ranges. Initially, the fine martensite is present even after tempering at 300 °C (Fig. 2b). However, the tempering at 600 °C may produce some bainite structure, which is observed by the morphological change of the microstructure (Fig. 2c). The recrystallization of ferrite grain and fine plate type M_2C type carbides may be

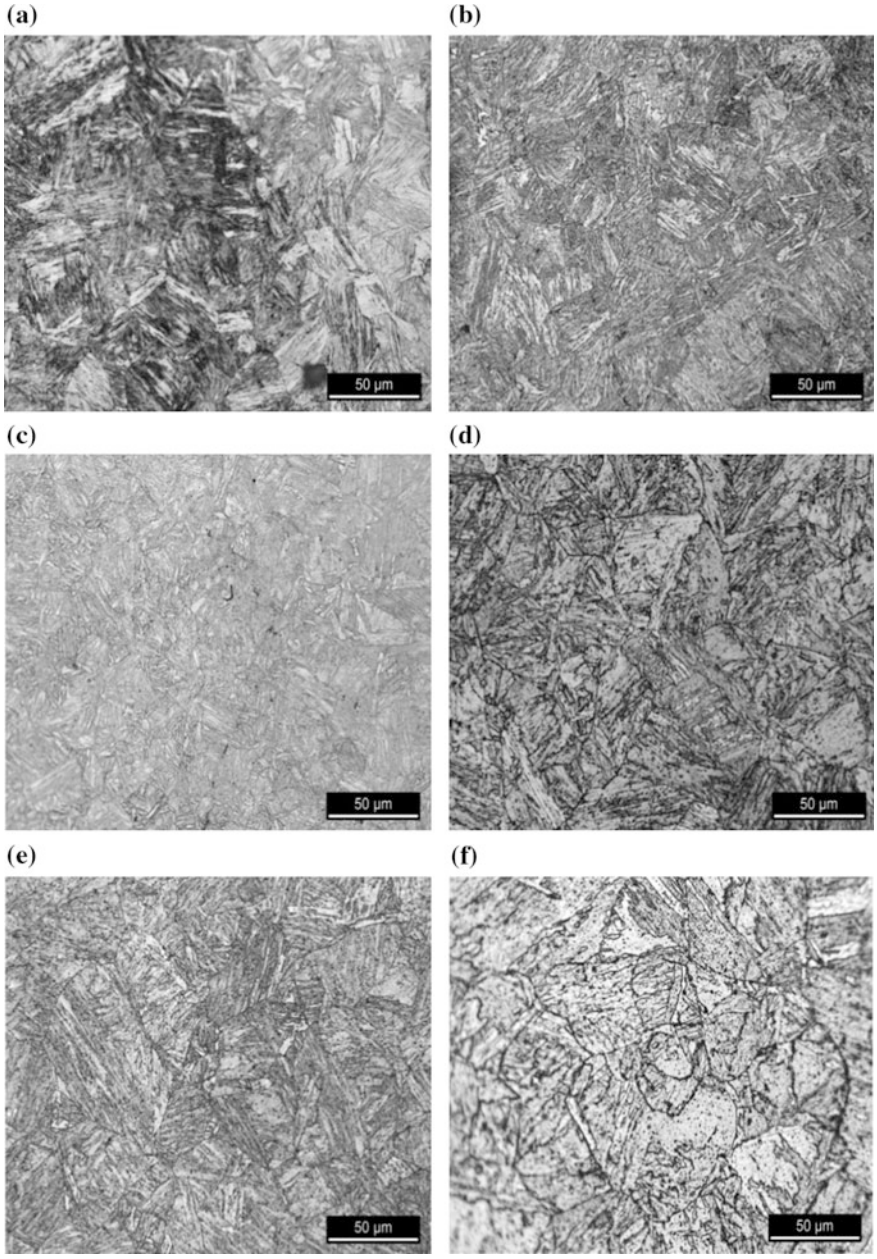


Fig. 2 Optical micrographs of SA213T22 sample **a** WQ, and tempered at **b** 300 °C, **c** 600 °C, **d** 700 °C, **e** 750 °C, **f** 800 °C

formed between 700 and 720 °C temperature ranges. For this reason, some big white color phases, possibly ferrite, are examined after tempering at 700 and 750 °C (Fig. 2d–e). At higher temperature (800 °C), the ferrite lath becomes coarsen and grain size is also increased (Fig. 2f). Moreover, the microstructural morphology of SA213 T22 steel is also the function of precipitates. Due to formation of different phases, the tempering treatment of SA213 steel may affect magnetic properties.

MBE parameter of tempered SA213T22: Figure 3a shows magnetic Barkhausen emission (MBE) voltage variation with tempering temperature. It is examined that that MBE voltage has not much changed after tempering at 600 °C compared to as-quenched state. Thereafter, the voltage increases and becomes maximum after tempering at 700 °C. It is already observed in optical micrographs shows that there is some ferrite nucleation at 700 °C. Simultaneously, the precipitation of M_2C type carbide also takes place between 700 and 720 °C. Owing to fine and densely distributed M_2C precipitates, the domain wall is not act as strongly pinned. Moreover, they also retard growth of ferrite grains. This all support for increase in Barkhausen emission. This demagnetizing field easily generate reverse spike domain at the boundary of these types of carbides which act as strong pinning for domain wall motion. Tempering at 800 °C completely changes the morphology of microstructure with the presence of larger size ferrite reduce effective area for domain wall motion, further reducing the MBE voltage.

MHL parameter of tempered SA213T22: Figure 3b shows the variation of coercivity with tempering temperature. It is clearly shows nonlinear variation of coercivity with tempering temperature. Initially due to fine martensite—bainite structure coercivity is very high. Upon progress of tempering, the coercivity decreases due to reduction in dislocation density. Further, the increase of tempering temperature, the ferrite nucleation causes the decrease of coercivity owing to the easy domain movement. Simultaneously, the precipitation of fine M_2C type carbide may restrict the domain movement at tempering temperature around 700–720 °C. It

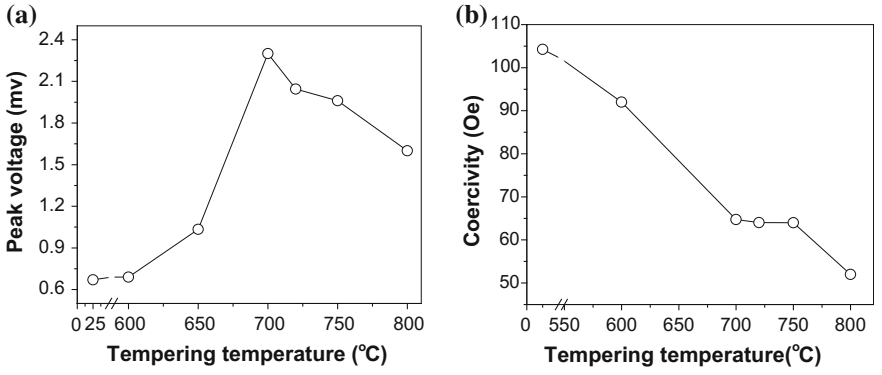


Fig. 3 Tempering temperature effect on magnetic parameters of **a** MBE **b** MHL

causes no change of coercivity in this temperature range. At higher temperature (800 °C) the dissolution of M_2C type carbide makes polygonization of ferrite resulting in the decrease of coercivity.

4.2 IF Steel

Microstructural evaluation: As received IF steel is cold rolled steel, describing elongated deformed grains aligned along rolling direction (Fig. 4a). Figure 4b–d show annealed microstructures of IF steel at different isochronal annealing temperature. Nucleation of strain free grains is revealed after annealing at 600 °C (Fig. 4b). Upon progress of annealing, number of strain free grains increases with the expense of deformed grains. This is called partially recrystallized stage, which is shown after annealing at 650 °C (Fig. 4c). Finally, the deformed grains are completely filled up with recrystallized grains after annealing at 700 °C (Fig. 4d).

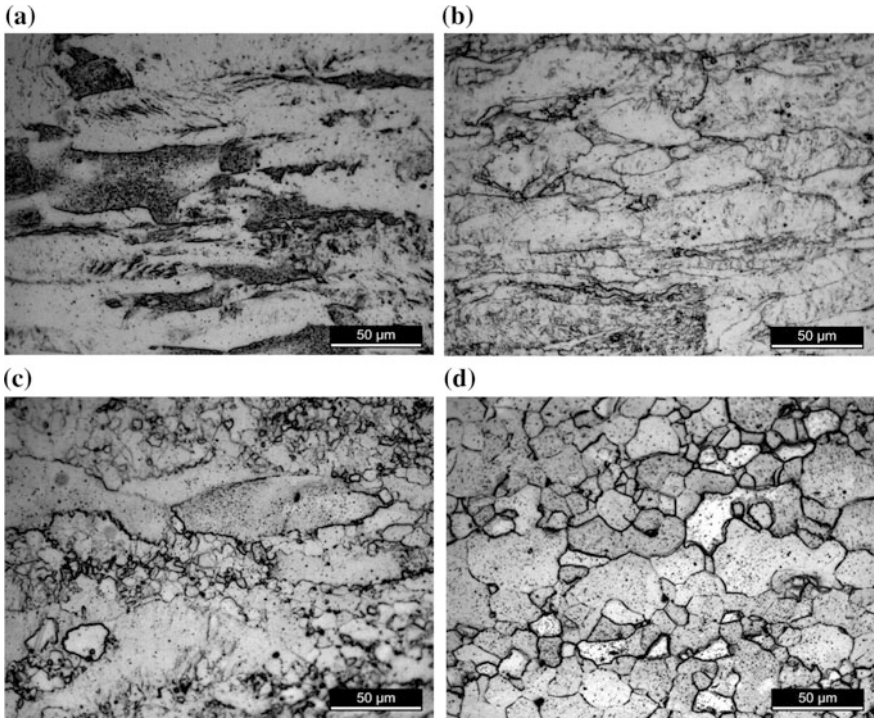


Fig. 4 Optical micrographs of IF steel sample **a** As-received, and annealed at **b** 600 °C, **c** 650 °C, **d** 700 °C

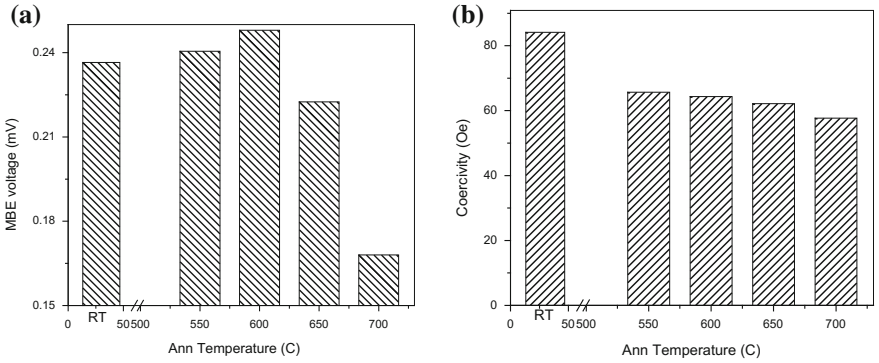


Fig. 5 Annealing temperature effect on magnetic parameters of **a** MBE **b** MHL

MBE parameter of annealed IF steel: Effect of annealing temperature on MBE voltage of cold rolled IF steel is represented in Fig. 5a. After annealing at 550 and 600 °C, MBE voltage increases slightly compared to as-received cold rolled steel at RT condition. Annealing above 600 °C, MBE voltage decreases with increasing annealing temperature. Till 600 °C, MBE voltage continuously increases due to the increase of mean free path of domain wall movement due to lowering of dislocation density. In contrast, above 600 °C, MBE voltage decreases owing to increasing the number of strain free grains, which causes the decrease of mean free path of domain wall movement.

MHL parameter of annealed IF steel: Coercivity of annealed IF steel is described in Fig. 5b. Unlike MBE voltage, coercivity progressively decreases within annealing temperature range. It happens because coercivity of ferromagnetic material is linearly related to dislocation density. Annealing treatment results in decrease of dislocation and generation of strain free grains. Therefore, magnetic sensor *MagStar* clearly explains the lowering of dislocation density during recovery and the generation of strain-free grains during recrystallization. These two stages (recovery and recrystallization) are very important for the processing of any steel in industry, which can be determined through destructive techniques like microscopy, hardness measurement, etc. The destructive techniques are not only time consuming but also demanding sophisticated instruments and highly skilled personnel. On the other hand, *MagStar* explain it rapidly with the simplest and easiest way.

5 Conclusions

Electromagnetic sensing device *MagStar* clearly describes microstructural phase changes in two different types steels, SA213T22 and IF steel. Since these two steels are used in different applications and produce completely different microstructures, *MagStar* has shown its credibility not only for different atmosphere but also in wide

applicability. Above all, *MagStar* is portable and cheap compare to its foreign counterpart of similar type instrument. Therefore, the device can be applicable for both industries and educational research.

References

1. Haldane, R.J., Yin, W., Strangwood, M., Peyton, A.J., Davis, C.L.: Multi-frequency electromagnetic sensor measuremen to ferrite/austenite phase fraction—experiment and theory. Scr. Mater **54**, 1761–1765 (2006)
2. Kumar H., Mohapatra J.N., Roy R. K., Justin Joseyphus R., Mitra A.: Evaluation of tempering behaviour in modified 9Cr–1Mo steel by magnetic non-destructive techniques. J. Mat. Proc. Tech. **210,** 669–674 (2010)
3. Gou, R., Zhang, Y., Xu, X., Sun, L., Yang, Y.: Residual stress measurement of new and in-service X70pipelines by X-ray diffraction method. NDT E Int. **44**, 387–393 (2011)
4. Kahrobaee, S., Kashefi, M.: Hardness profile plotting using multi-frequency multi-output electromagnetic sensor. NDT E Int. **44**, 335–338 (2011)
5. Yin W., Karimian N., Liu J., Hao X.J., Zhou L., Peyton A .J., Strangwood M., Davis C.L.: Measurement of electromagnetic properties of power station steels. NDT E Int. **51**, 135–141 (2012)
6. Roy, R.K., Panda, A.K., Mitra, A. In: Proceeding of 6th International Conference on Sensing Technology (ICST 2012), Kolkata, India, 18–21 Dec 2012, pp. 226–229. IEEE Xplore
7. Barkhausen, H.: Phys. Zeitschrift **20**, 201 (1919)
8. Gateier-Rothea, C., Chicois, J., Fougeres, R., Fleisch-mann, P.: Acta Mater. **46**, 4 (1998)

Developing a Smart Navigator for Surveillance in Unmanned Zones

Pooja Nag, Sumit Shinde and Kapil Sadani

Abstract The current work reports of an obstacle avoidance and object tracking algorithm integrated and tested on robot. The system is designed with high torque geared DC motors and it has a payload capacity of up to 3 kg. The robot is fully autonomous and it is designed to patrol high security/hazardous zones and dynamically report any suspicious activity which when observed initiates an alarm to the security for further action whiles it continues to track and report the suspect. This has a major advantage over conventional CCTV systems in terms of cost and memory requirements and would not require constant human supervision. The robot proposed hereof uses wheel-encoders, ICbased Gyroscope, IR Line Laser, and spy camera as the basic sensing elements. It also has a smart charging feature which makes it energy efficient.

Keywords Object tracking · Obstacle avoidance · Surveillance

1 Introduction

Robots are commonly used in environments where human presence is either risky or redundant. Risks, such as high pressure, temperature, or corrosive atmospheres and repeated tasks such as quality inspection of finished goods especially food products. The robots used in such conditions are generally telecontrolled with live data steaming and annunciation protocols [1]. Reports of developing a telecontrolled robot for live health monitoring in domestic environments. A similar tele-

P. Nag (✉)
Department of Mechatronics, Manipal Institute of Technology, Manipal University, Manipal 576104, Karnataka, India
e-mail: pooja.nag@manipal.edu

S. Shinde · K. Sadani
Department of Instrumentation and Control, Manipal Institute of Technology, Manipal University, Manipal 576104, Karnataka, India
e-mail: sadani.kapil@manipal.edu

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_25

controlled robot is described in [2] which is focused on the security tasks over a multi-threading platform. On the other hand, there are the autonomous robots which are based on preprogrammed tasks or on complex agent architectures. These robots are generally even more difficult to develop but they are free from constant human assistance. An autonomous explorer robot was presented in [3–6] which is capable to navigate between locations in an intra-urban environment. Mobile robots are also commonly used for measurement and recon operations in human-risky environments such as operations under toxic air conditions. In [7] is presented an autonomous robot agent for searching gas/odor sources in an indoor environment using a biologically inspired algorithm.

In this work, we report of an autonomous low-cost robot which can patrol unmanned zones and send live video in case a suspicious activity is sensed. The application varies from surveillance to fire to gas leaks. This helps reduce the power and data storage overhead required for existing CCTV-based surveillance systems.

2 Experimental

2.1 Apparatus

In this work, 1× Arduino Mega 2560 which continuously communicates with the Computer serially, 3× SR-04 Ultrasonic sensors, 1× Laser emitter with wide angle lens, 1× L293D Motor driver, 2× High torque geared DC Motors, 2× Wheels, 2× wheel-encoder, 1× castor wheel, IC-based Gyroscope, ADXL Accelerometer, MQ-2 Gas sensor, 2× vision cameras, 1× robot vehicle chassis have been used. LabVIEW Machine-vision toolbox for Video and Image processing. Algorithms are also developed on LabVIEW platform for image acquisition and motor controller (Fig. 1).

2.2 Features

(a) Obstacle detection and visualizing its size, shape position. After reading the surrounding objects the shortest and safest path is chosen by robot avoidance using ultrasonic sensors

A 3-way arrangement of ultrasonic sensors facing left, forward, and middle is used for obstacle detection and avoidance. Ultrasonic sensors work on the principle of calculating distance based on the time taken in travel from emitter to receiver in the sensor.

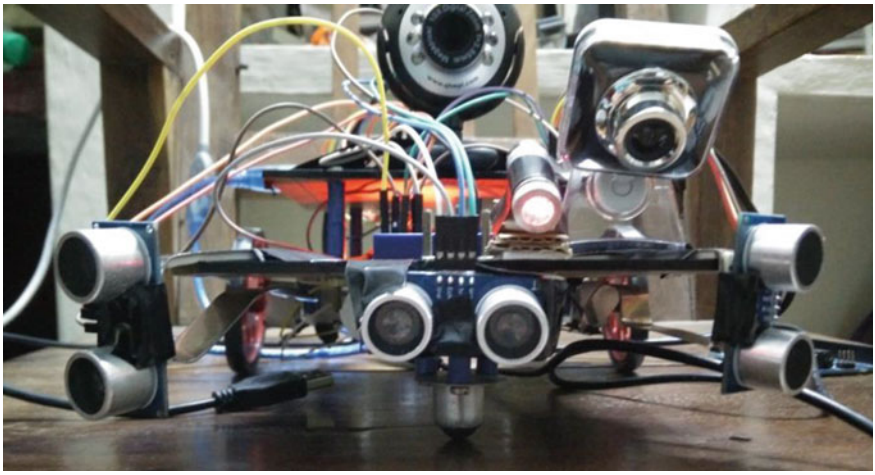


Fig. 1 The designed Robot

The movement of motors is chosen in such a way that the ultrasonic sensors help to maintain a safe distance of 15 cm from an obstacle.

The above code snippet depicts the readout from the individual ultrasonic sensors and the subsequent logic to control the motors. A local variable named Obstacle is taken from another parallel running code to further improve the movement of robot during obstacle detection and avoidance. Ultrasonic sensors are chosen because of no inherent lags in their response.

(b) Using camera and laser line projection

The following figures illustrate the basic principle of a 2D method of object detection using a 1D line as shown in Fig. 2a, b.

This technique yields better obstacle detection when used in parallel with ultrasonic sensors. In such a setup, three ultrasonic sensors facing left, center, and right are used. If the left or right ultrasonic sensors detect an obstacle, the robot stays true to its path. If the center sensor along with either left or right sensor detects an obstacle, the robot turns toward the opposite direction respectively. Here, the laser line helps in detecting obstacles in the blind spots between the individual ultrasonic sensors.

The laser line at times will be unrecognizable in some backgrounds when direct pattern matching algorithms are used on the image. To aid in correct recognition of the laser line, the scene has to be enhanced in contrast to better perceive the sharp change in brightness from the laser line with respect to its background (Fig. 3).

Here, the video feed acquired by the camera is processed continuously in LabVIEW software. Video feed is converted into a sequence of images and each image is smoothed at first stage. To characterize the image at each key location, Image I is processed at each level of pyramid to extract image gradients and

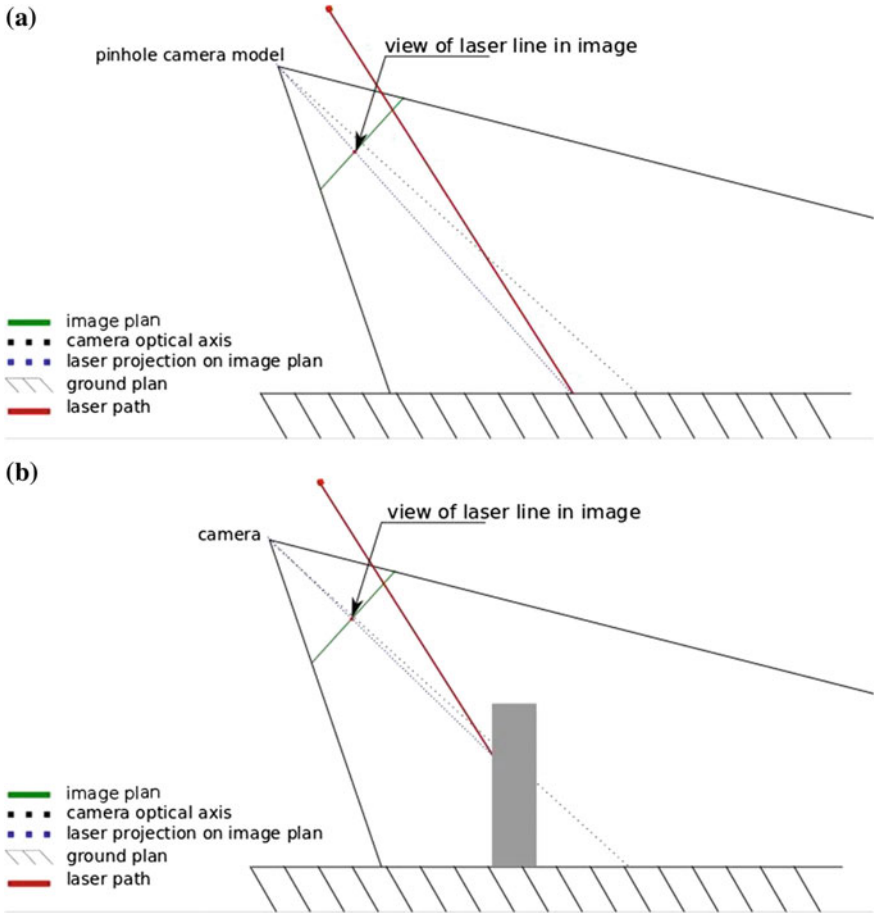


Fig. 2 **a** No obstacle scenario gives pattern matching score maximum. **b** Obstacle detected as the score of pattern matching is minimum

orientations. At each pixel Iij the image gradient magnitude Mij and orientation Rij are computed using pixel differences.

$$M_{ij} = \sqrt{(I_{ij} - I_{i+1,j})^2 + (I_{ij} - I_{i,j+1})^2}$$
$$R_{ij} = \text{a } tan^2(I_{ij} - I_{i+1,j}, I_{ij} - I_{i,j+1})$$

The Fig. 4 describe the working of laser line-based obstacle avoidance using direct patter-matching algorithm on the grayscale image. When the laser line is distorted the template match fails and thus, it is inferred as presence of an obstacle. At this point a suitable signal is given to the motors so that the robot is able to avoid the obstacle (Fig. 5).

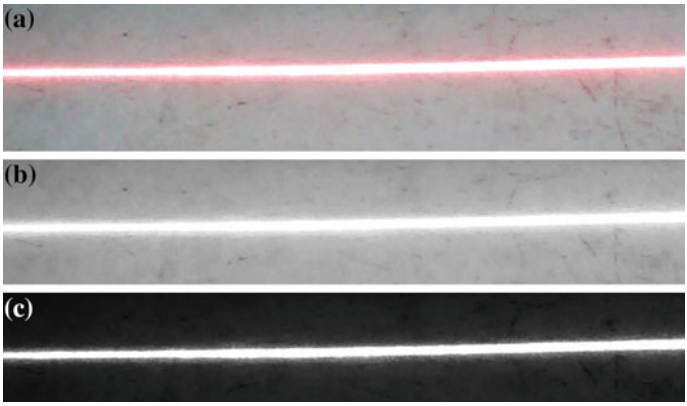


Fig. 3 **a** Laser line image. **b** Grayscale conversion of above image. **c** Image enhancement using exponential lookup table

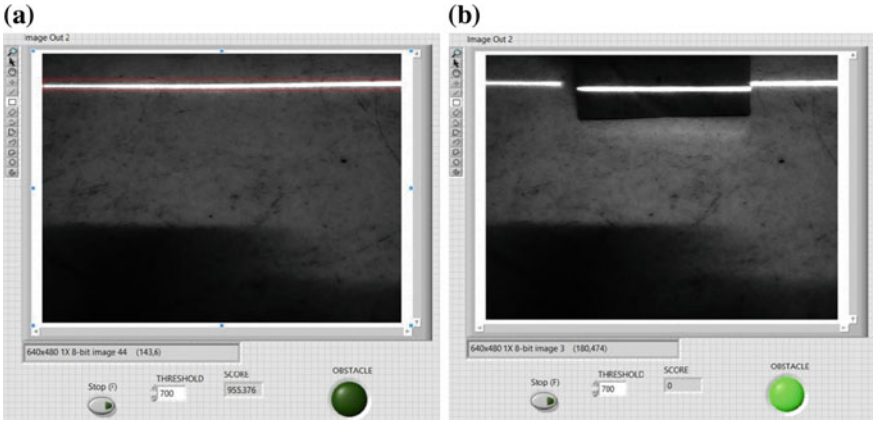


Fig. 4 **a** Non-distorted laser line implying absence of obstacle. **b** Distorted laser line implying presence of obstacle

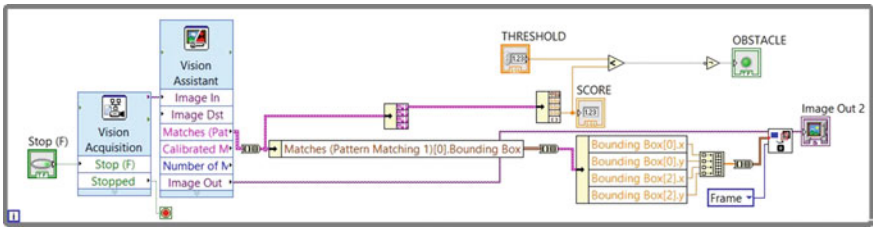


Fig. 5 VI for obstacle detection using Laser

The above code snippet is used for detecting any change in the image of a laser line projected in front of the robot. The working principle is based on the parallax error to perceive a visible change in laser line when an object come in contact with the laser. The camera is made to face downward from an elevation, pointing at a flat surface. The laser line projector is made to point at the flat surface, from a location below the camera. The axes of camera and laser should not be parallel and should be adjusted as long as the laser line is visible in the camera's field of vision. In this case, the line is 15 cm away from the camera.

To a camera, the laser line will appear as a mix of red and white due to the intensity of projected laser beam. A subsequent grayscale image conversion helps to simplify the processing speed by removing the color plane in the image. An exponential lookup table operation is applied to the image in order to enhance the contrast in the image, so as to emphasize on the differences between laser line and background (Figs. 6 and 7).

When there is a cliff, the laser disappears from the camera field of view and a random image pattern is observed. This pattern has little correlation with the laser line, and as a consequence an obstacle is acknowledged, making the robot stop.

(c) Smoke and Gaseous impurity detection

Chemoresistive gas sensor for detection of sensitive for Methane, Butane, LPG, smoke is used. This sensor is sensitive for flammable and combustible gasses. The gas sensors commercially available (and used here) generally comprise a thin film transition metal oxide as the sensing element. The thin film surface has highly

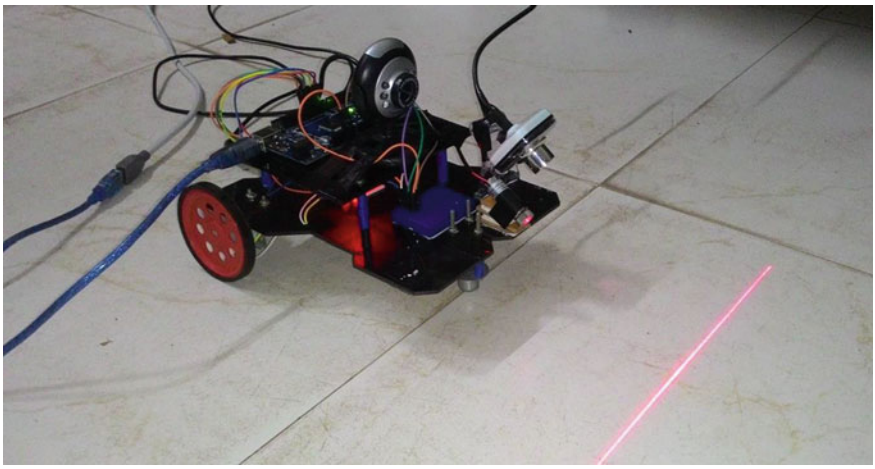


Fig. 6 The Robot scanning through its path and selects the shortest and safest path. The laser line at times will be unrecognizable in some backgrounds when direct pattern matching algorithms are used on the image. To aid in correct recognition of the laser line, the scene has to be enhanced in contrast to better perceive the sharp change in brightness from the laser line with respect to its background

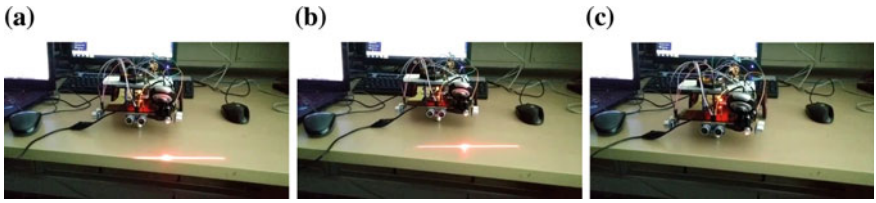


Fig. 7 **a** Robot analyzing the plane surface with laser beam. **b** Robot detects the cliff. **c** Robot control the wheels to moves in safe position

reactive oxygen atoms on its surface; when a reducing gas (for which the sensor surface properties have been tuned) comes in contact, it reacts with the surface oxygen in a heterolytic dissociation process and the oxygen atoms are carried off the surface releasing the electrons to which it was bonded. Thus, some electrons are released leading to an increase in the conductance which may be used for annunciation in conjunction with a signal conditioning circuit. In this case, since it is a 10 bit ADC in the processor unit, the analog readout from the gas sensor is converted to bits accordingly to accurately measure the conversion result. The following is a code snippet depicting the gas sensor's output used to trigger the alarm.

The role of the sensor is to check for presence of any smoke or gaseous impurity in the environment and trigger an alarm. This alarm will then be subject to acknowledgement by the operator who will then decide whether to acknowledge the alarm or not. If it is acknowledged, a video feed of the surrounding is captured and uploaded on the cloud (internet) or until the operator desires so (Figs. 8, 9 and 10).

To characterize the image at each key location, the smoothed image A at each level of the pyramid is processed to extract image gradients and orientations. At each pixel, Aij, the image gradient magnitude, Mij, and orientation, Rij, are computed using pixel differences.

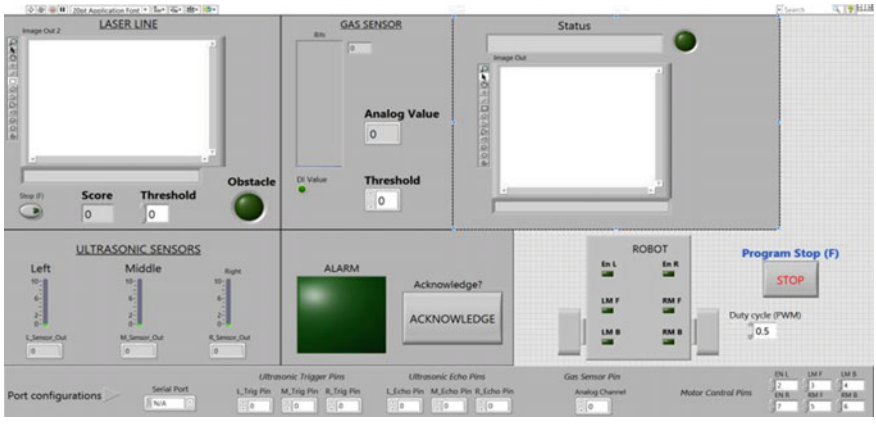


Fig. 8 Data Acquisition and annunciation from the gas sensor

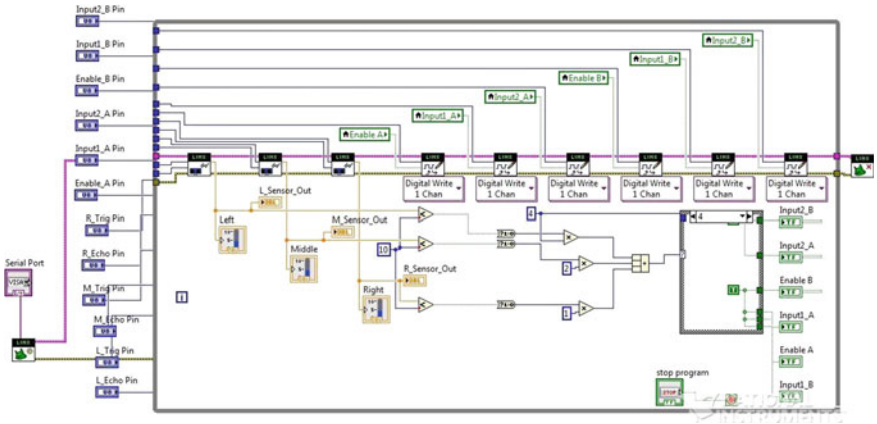


Fig. 9 VI for data acquisition from the ultrasonic sensors

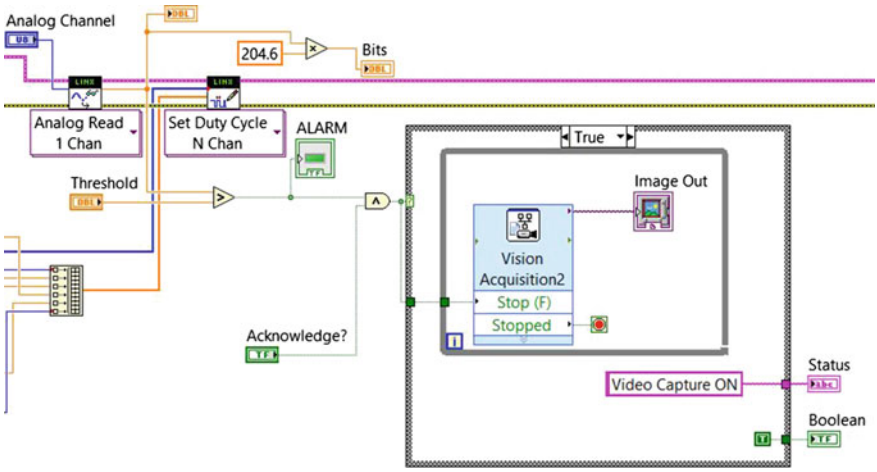


Fig. 10 The Graphical User Interface for the designed Robot in LabView

3 Conclusion

The developed Robot is capable to move through given path autonomously. We have used IC-based advanced gas sensors and actuators thus making the product affordable. The algorithm reported hereof has negligible time lag and is very effective for surveillance applications. An extended application of such robots may be in kitchens of restaurants to check for gas leaks in the absence of chefs and workmen. Also such robots would find great application in parking lot surveillance in big housing complexes. This not only provides the low-cost solution in many security applications but also gives user a scope to monitor the surrounding

remotely. This simple device if equipped with suitable gas sensors and IP-based cameras may as well be used for detecting and monitoring allowable levels of noxious gases in certain petrochemical industries. The scope and application of the robot is really vast and diverse.

References

1. Labonte, D., Michaud, F., Boissy, P., Corriveau, H., Cloutier, R., Roux, M.A.: A pilot study on teleoperated mobile robots in home environments. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 4466–4471, October 2006
2. Birk, A., Kenn, H.: Roboguard, a teleoperated mobile security robot. In: Control Engineering Practice, vol. 10, no. 11, pp. 1259–1264. Elsevier (2002)
3. Lidoris, G., Rohrmuller, F., Wollherr, D., Buss, M.: The Autonomous City Explorer (ACE) project—mobile robot navigation in highly populated urban environments. In: IEEE International Conference on Robotics and Automation (ICRA), pp. 1416–1422, May 2009
4. Biswas, J., Veloso, M.: Wifi localization and navigation for autonomous indoor mobile robots. In: IEEE International Conference on Robotics and Automation (ICRA), pp. 4379–4384, May 2010
5. Maier, D., Kleiner, A.: Improved GPS sensor model for mobile robots in urban terrain. In: IEEE International Conference on Robotics and Automation (ICRA), pp. 4385–4390, May 2010
6. Lee, K., Chung, W., Yoo, K.: Kinematic parameter calibration of a carlike mobile robot to improve odometry accuracy. Mechatronics **20**(5) (2010)
7. Ferri, G., Caselli, E., Mattoli, V., Mondini, A., Mazzolai, B., Dario, P.: SPIRAL: a novel biologically-inspired algorithm for gas/odor source localization in an indoor environment with no strong airflow. Robot. Auton. Syst. **57**(4), 393–402 (2009)

A Simple Flow Measurement System for Rotameters Using Webcam

Pooja Nag, Sumit Shinde, Dayananda Nayak and Kapil Sadani

Abstract In this work, we report of a noncontact level measurement based on machine vision techniques. Data has been acquired with a high-definition webcam @ 30 frames per second. Image-processing algorithms have been developed and optimized using LabVIEW. The measurement may then automate a servo actuated hand-valve. On testing, the system was found to be linear, repeatable, and accurate to $\pm 0.5\%$ of the measured value for 50 measurements carried out. In the scheme are proposed hereof, the levels may just be remotely observed on a LabVIEW SCADA panel and used as a set point to a regulatory control-actuator system. The system is rugged and can be calibrated to 0.05% precision for all rotameter types used in any environment from a simple DM plant to a petrochemical industry for a fixed camera distance.

Keywords Computer vision · PID controller · Pattern matching · Template matching

1 Introduction

Flow measurement is of utmost importance in applications ranging from municipal water corporations, process industries to common households. The Rotameter is one of the cheapest and most reliable mechanical instrument used for measuring flow of clean homogeneous fluids. The position of the float is calibrated onto a

P. Nag
Department of Mechatronics, Manipal Institute of Technology,
Manipal University, Manipal, India

D. Nayak · K. Sadani (✉)
Department of Instrumentation and Control, Manipal Institute of Technology,
Manipal University, Manipal, India
e-mail: sadani.kapil@manipal.edu

S. Shinde
Department of Embedded Systems Engineering, Albert Ludwig University of Freiburg,
Freiburg im Breisgau, Germany

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_26

glass/mica scale calibrated in terms of flow. Generally, the operator observes the float position and fixes a valve position for the desired flow rate.

1.1 Image Processing

Recent advances in image processing has tremendous impact on society ranging from food quality monitoring [1, 2] medical diagnostics [3] surveillance [4–6] and so on.

This work is focused on automating this process by developing a machine vision-based control system that may be used for continuous monitoring the flow rate in rotameters. The system is trained to visually read the position of the spinning float in rotameter and track the same while in operation. A closed water flow circuit powered by a low-power centrifugal pump is used for the purpose; the computer reads the float position through a web camera which is calibrated in terms of flow.

1.2 Template Matching Technique

Template matching is a high-level computer vision technique in which parts of an image are constantly compared with a predefined template denoted as the area of interest. The algorithm should be so defined such that the match is computed irrespective of the orientation of occurrence and the local illuminance of the surroundings.

They mainly need two things to be defined one is template image which is the reference image of an object and source image or input image. Here, we want to identify all input image locations at which the object from the template image is present. It is expected that not to identify the scaled or rotated occurrences and the algorithm developed takes care of the same.

In LabView, one just needs to define a template, which is essentially a part of the frame which is to be tracked, in our case, the float. The top of the float is defined as the region of interest and its coordinates are linearly related to the position which in turn is linearly related to the flow rate.

2 Experimental

Figure 1 illustrates the experimental setup where this work has been carried out. The camera used is a simple webcam purchased from Amazon.in. It is installed at a prefixed position on a custom made stand; the camera height and tilt being adjustable. The figure to the top is that of a standard flow trainer kit with water as the process fluid. Water is pumped in with a controllable pump from the reservoir to the Rotameter. A hand-valve attached to the input of the rotameter. The objective is to vary the pumping rates and control the position of the hand-valve so as to keep the flow rate

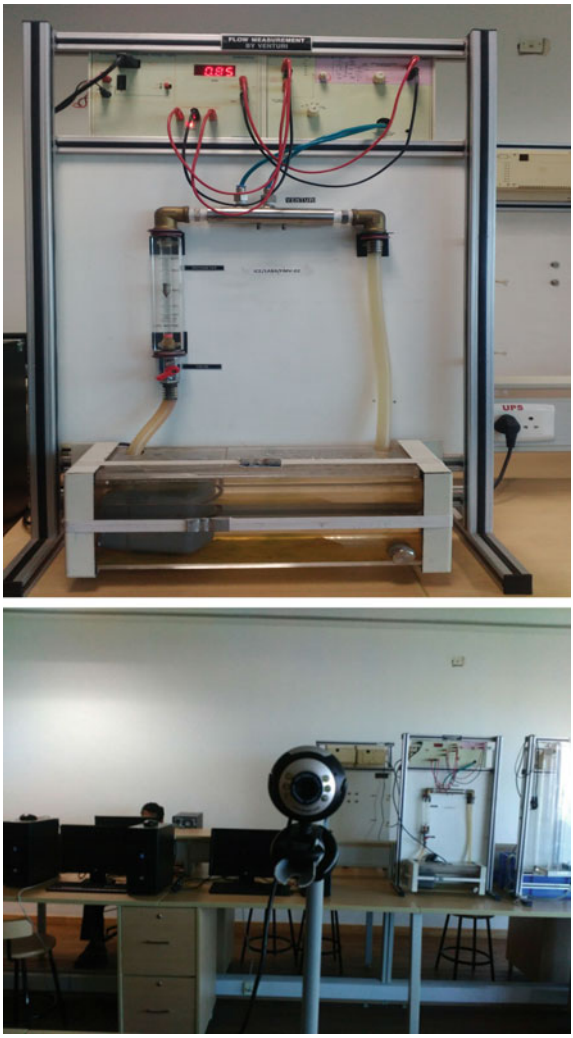


Fig. 1 The experimental setup

fixed. For the purpose, first, we devise a strategy to measure the flowrate with the webcam. The graphical program used to acquire the image is shown in Fig. 2.

The code illustrates a very simple vision acquisition technique where the camera continuously acquires images of the rotameter. A template, i.e., the top of the float has been defined as the region of interest and the algorithm tracks the coordinates (position) of this template. In a rotameter, for the measurement of liquid flows, generally as the flow rate increases, the float moves upwards and as it reduces, it moves downwards. The principle of variable area flowmeters defines the relation between the position of the float and the flow to be being linear. Hence, if the coordinates of the float be measured, it can be mapped to a linear fit to determine the flow rate (Fig. 3 and Table 1).

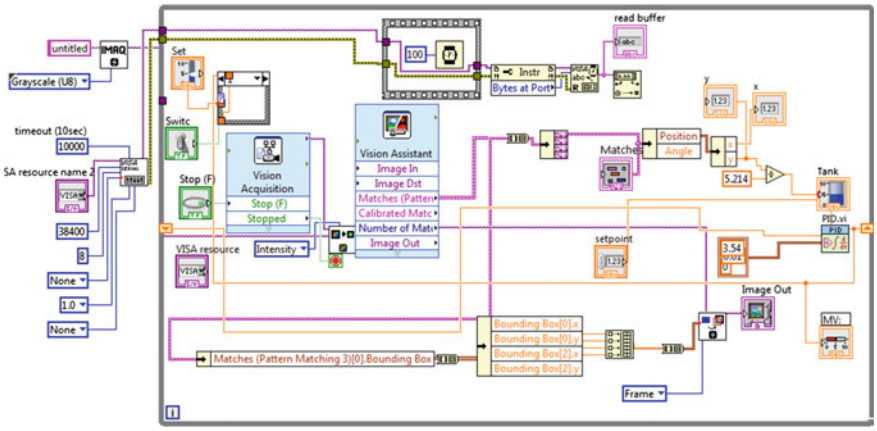


Fig. 2 Image acquisition VI

Fig. 3 Rotameter float is at different positions sensed by camera and calibrated in terms of level

Table 1 The following table relates the position of the float with the flow as measured by the webcam and as measured visually

Actual flow rate (in LPH)	Rotameter-float position (cm)	Pattern co-ordinates (X, Y)	Calculated flow (in LPH)	Error (in LPH)
50	4	(218,202)	50.04	0.04
75	6	(218,304)	75.10	0.10
100	8	(220,401)	100.02	0.02
125	10	(221,505)	125.14	0.014
150	12	(219,603)	150.08	0.08
175	14	(221,702)	175.06	0.06
200	16	(220,804)	200.10	0.10

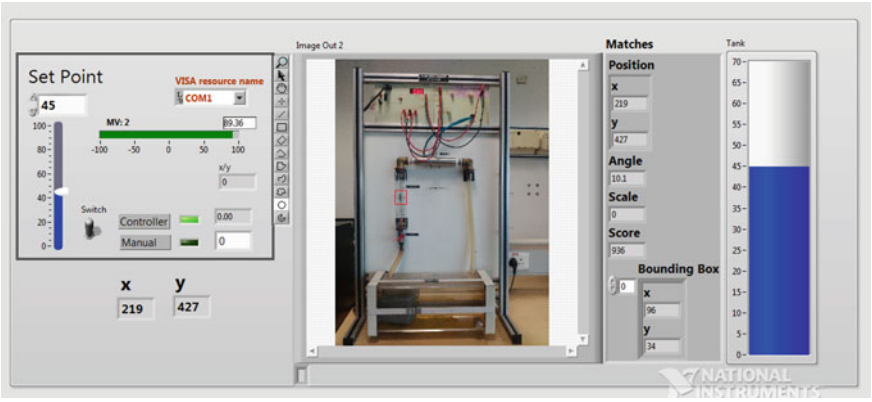


Fig. 4 The SCADA for the noncontact flow measurement system

It is thus very clear that this technique of noncontact flow measurement is very simple and rugged. The coordinates may be stored in a lookup table synonyms to specified flow rates. With any suitable flow as the set point a PI or a PID controller may be suitably designed to control the flow rate. The choice of the algorithm depends on the speed of response desired. If the system has a large capacity and offset is acceptable a simple linear proportional controller would serve the purpose.

Shown below is the front panel of the acquisition system designed on LabView. The front panel is in the form of a supervisory control as depicted in Fig. 4.

3 Conclusion

A very simple yet a very elegant scheme of noncontact level measurement has been discussed. The system is cost effective as the only requirement is that of a webcam and a local FPGA processor supporting LabView. The system can be incorporated with minor modifications in process industries; such as a simple fan not allowing dust to segregate on the camera assuming proper lighting available. Background subtraction algorithms may be suitably incorporated to make the system more efficient. Proper lighting if not ensured may lead to the system failure. This technique has vast applications in DM plants, where the fluid is generally clean and light is abundant. The system designed can control flow rate by actuating a servomotor to regulate a valve. However, the accuracy of the system is limited by the Rotameter itself which suffers from inherent errors. However, this approach is elegant in the sense that constant human supervision for simple controls can be avoided.

References

1. Yao, A., Gall, J., Gool, L.V.: Coupled action recognition and pose estimation from multiple views. Int. J. Comput. Vis. 16–37 (2012)
2. Barbulescu, A., Gong, W., Gonzalez, J., Moeslund, T.B., Roca, F.X.: 3D human pose estimation using 2D body part detectors. In: 21st International Conference on Pattern Recognition, pp. 2484–2487 (2012)
3. Cubero, S., et al.: Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables. Food Bioprocess Technol. **4**(4), 487–504 (2011)
4. Biswas, J., Veloso, M.: Wifi localization and navigation for autonomous indoor mobile robots. In: IEEE International Conference on Robotics and Automation (ICRA), pp. 4379–4384, May 2010
5. Maier, D., Kleiner, A.: Improved GPS sensor model for mobile robots in urban terrain. In: IEEE International Conference on Robotics and Automation (ICRA), pp. 4385–4390, May 2010
6. Lee, K., Chung, W., Yoo, K.: Kinematic parameter calibration of a car like mobile robot to improve odometry accuracy. Mechatronics **20**(5), 582–595 (2010)

Sensor Search Using Clustering Technique in a Massive IoT Environment

Nandhakumar Ramachandran, Varalakshmi Perumal, Sakithya Gopinath and Monika Jothi

Abstract The Internet of things contains billions of interconnected things such as physical objects, animals, or human beings that give the power to transfer information over a network without human interaction. The huge amount of data continuously generated by these sensing devices raises the challenge of searching for the most relevant sensor data according to user's query. Also, the queries specified by users are in natural human language which cannot be processed by sensors. To handle these challenges, devices in an IoT environment are grouped together to form clusters which reduces the search space. Every device is given a unique ipv6 address to identify it within the network. Then, the user's abstract query is transformed into low level form which can be recognized by sensors. Experimental results show that the search time and response time is improved using this approach in large scale IoT environments.

Keywords IoT · Sensor search · Query processing · IPv6

1 Introduction

The Internet of things contains group of interrelated devices that enable them to exchange data. It brings internet to the physical world so that the objects can be tracked and managed remotely. A thing, within the Internet of Things, can be

N. Ramachandran (✉) · V. Perumal · S. Gopinath · M. Jothi
Department of Computer Technology, MIT Campus, Anna University, Chennai, India
e-mail: nandhakumarr03@gmail.com

V. Perumal
e-mail: varanip@annauniv.edu

S. Gopinath
e-mail: sakithya.gopinath@gmail.com

M. Jothi
e-mail: monika.jothi@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_27

someone with a cardiac monitor implant, an automobile that has inbuilt sensors which will alert the driver when tire pressure is low or any other natural or man-made object that may be allotted an IP address. The main aim of IoT is to create a better world for people, where things around us communicate with one another and know what we like, what we want, and what we need and act accordingly without specific instructions [1–3].

IoT consists of three layers, such as application layer, network layer, and sensing layer. The sensing layer consists of various sensing devices, such as controllers, sensor devices (smoke, temperature, pressure, light, and humidity), RFID tags, camera, energy meter, and actuators etc. which have limited energy, memory and processing power. In monitoring and detection applications, it is important to select the specific sensors that cause a particular effect. Our objective is to help users retrieve the most relevant information according to their needs.

In this paper, we propose a cluster-based searching approach which can be applied to any IoT application. As IoT consists of a huge number of sensors, the energy of these sensors will deprive quickly if they start sending data directly. Therefore, the sensors are grouped together into clusters based on their energy level and proximity. These sensors are assigned address in a hierarchical manner to identify them uniquely in a large-scale network [4]. We process the user query which is in natural human language and transform them into form recognized by the sensors using predefined database of keywords. Then this query is matched with the sensors to perform the search and the data is sent back to the user.

In the remainder of the paper, we review the literature and discuss the existing related works in Sect. 2. The proposed algorithm is presented in Sect. 3. Experimental results and discussion are presented in Sect. 4. Section 5 concludes the paper.

2 Literature Survey

The ideal solution must work in such a way that after the user enters a query, results should be given to the user without making the user selecting the specific sensors from which they require the information. There are various approaches designed for this, such as context-based search and content-based search. The query can also be given in any standard languages, such as SPARQL, SQL, and POSTGRESQL etc.

A context-aware sensor search, selection, and ranking model, called CASSARAM [5] efficiently chooses the required sensor from the huge group of sensors present based on user needs. It is a context-based search based on the sensor properties, such as location, battery life, dependability, etc. When the number of sensors increases, the performance of the system decreases. Also the storage requirement is directly proportional to the number of context properties of each sensor considered.

A primitive method called sensor ranking [6] is a content-based sensor search which matches the user's query based on the estimated probability. The basic idea is to create a list of sensors which are ranked based on the content. The sensor with highest rank has more probability to match the user query. A fuzzy-based sensor search [7] in the web of things is based on similar data outputs in the past. Fuzzy sets are used to represent the past output computed by each sensor and are indexed in the database. Given the output of one sensor, for each indexed sensor, similarity scores are calculated by using sensor range difference.

Ding et al. [8] proposed a multimodal search on data based on spatialtemporal, value-based, and keyword-based(SVK) search. The SVK search engine uses, full text keyword B+ tree index for keyword-based search, IoT spatial temporal R Tree index for spatialtemporal-based search and value symbolized keyword B+ tree index for value-based search. An effective context-aware method to cluster sensors is proposed in the form of sensor semantic overlay networks (SSONs) [9], where the similar category sensors are grouped. Sensors are clustered in each SSONs by context-aware ant-based clustering algorithm. The iterations of the AntClust algorithm increase in the linear way as the sensor count increases. The accuracy of search result is very less when compared with CASSARAM.

A novel ontology-based method is stated in [10] where the abstract queries are hierarchically decomposed into many simple queries in a form understood by sensors. Ontology library is defined either by human or by intelligent ontology building applications. Effectiveness evaluation manager calculates the efficiency of each subqueries and return the answer to the users. As there as huge amount of sensors deployed around the place, it is difficult to preload all the data in the database.

3 Proposed Work

3.1 High-Level Model Overview

Our objective is to allow users to search and retrieve information from sensors according to their need. As the physical layer of IoT consists of numerous sensors as shown in Fig. 1, the first step is to form clusters of these sensor nodes based on the coverage range and then assign unique identifiers to the nodes in a hierarchical fashion. The user enters the query at the application layer which is then decomposed and prioritized based on the set of predefined rules. Then the prioritized subquery is transformed into a low -evel form and then sent to the base station as given in Algorithm 1. The base station will perform depth first search by identifying the specific controllers and the sensor nodes. The data is retrieved from the sensor and sent back to the user.

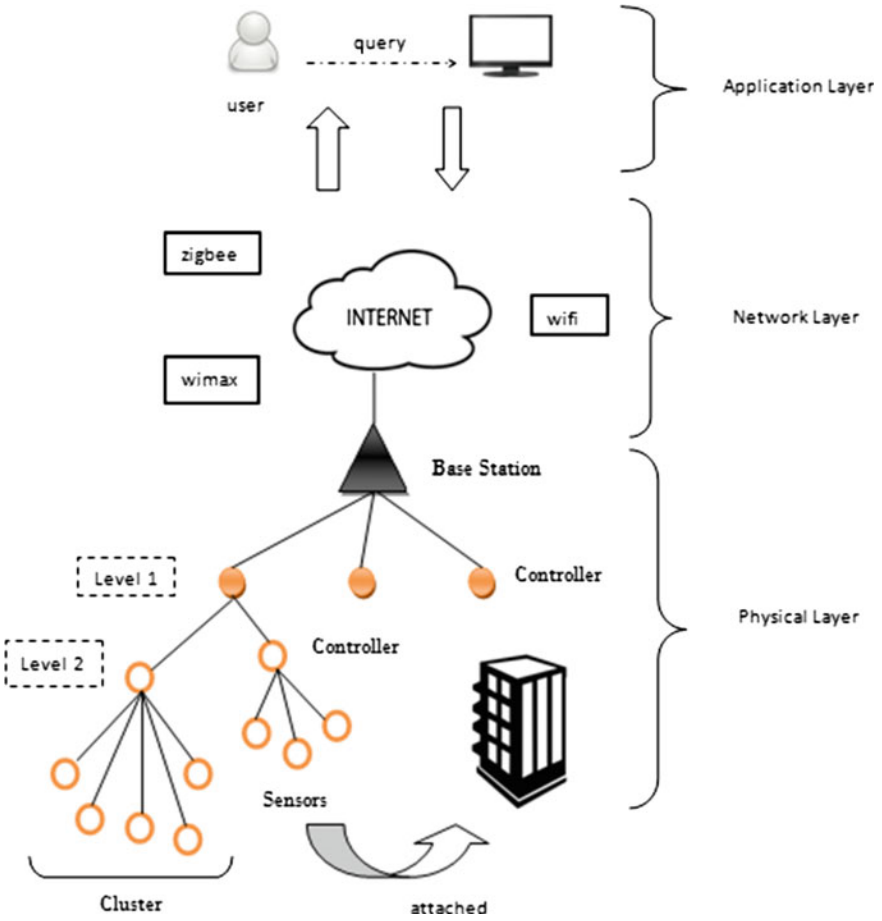


Fig. 1 Architecture diagram

Algorithm 1

```
Input:  Query from user, Set of clustered sensors
        table - keywords and  predefined set of rules.
Output: Data requested by user.
   1: Decompose the given query into words.
   2: Check for potential keywords
   3: Assign priorities to each word based on table
   4: Create a table in priority and corresponding identifier
   5: Send it to base station.
   6: Search the sensors.
   7: Return value to the user
```

3.2 Cluster Formation

Base station:

It is responsible for collecting the required information from the specified nodes and transferring it to the higher layers. It initiates communication by sending hello packets to its neighbors. It stores information about the controllers that are within its one hop distance. We assume that the base station is located at the centre of the surface area provided. The processed query is sent to the base station which will initiate the searching process.

Controller:

They gather information from the nodes within its cluster. Controllers are elected based on its proximity to the neighbors and its distance from the base station. It also performs data aggregation. It stores the information about the nodes that are within its control. It is a normal sensing node with specialized functions.

Sensing devices:

They have their own unique id. They are tiny, low powered, battery operated devices. They sense the environment and gather the data. The various types of devices present here are temperature, pressure, humidity, heat, light sensors, RFID tags, actuators, etc.

Hop count:

The hop count refers to the number of intermediate devices through which data must pass between source and destination. The BS will set its hop count as 1 and broadcast it to its one hop neighbors. These nodes will increment the count by one and send it to its one hop neighbors.

Neighbor count:

Every node will calculate its number of neighbors by broadcasting a hello message consisting of its count along with its id to all the neighbors.

First, every device must register its hardware address with the base station so that we can know what kinds of devices are available in our specific environment. Since the proposed application will not be present in any remote area, the sensors can be replaced as and when the battery level drops. Therefore, we don't consider energy as a major factor while forming clusters. The base station will initiate the communication by sending broadcast message to the devices within its range. It will elect the first level of controllers based on the region the controller is present so that, each region will have a controller which can communicate with the base station. Also unique identifiers are assigned to each device as and when the clusters are formed. If two nodes have the same neighbor count, then the node whose distance is far away from the base station, but still reachable to the controller is selected.

3.3 Device Classification

There are large numbers of devices available in IoT environment; therefore it is very difficult to identify a specific device within a network. These devices are assigned a unique ipv6 address format as shown in Table 1. The 64-bit address is divided into many parts where each part will help us identify what type of device is present and the location it is present.

While forming clusters, we assign identifiers to the devices simultaneously in a hierarchical manner. Initially, all the bits are set to 0. Based on the predefined values, the base station's ID will be assigned to an 8 bit number. Then for every device, the number of hops from the base station is calculated and assigned to level ID. Then, the identifier of controller of the device is specified. Also, as there are many different kinds of devices (temperature, light sensors) present it IoT, it is essential to specify the device type. The value ID is used to denote the specific device number within that cluster. Each unique Identifier is obtain based on the BS ID, Level ID and Controller ID.

3.4 Query Decomposition

The user will enter the query in complex and abstract natural English language. We need to discover related sensors in order to respond to the query. We assume that a query is a combination of keywords, attributes and the activity that is to be performed on the given set of sensors. Thus a query is defined as, {K1, K2, …, Kn; Attr1, Attr2, .., Attrn; A1, A2, …, An} where Ki is a keyword, Attri is an attribute and Ai is an activity as given in Algorithm 2. Any given keyword in the query is mapped with a corresponding ontology defined which contains the keyword, its priority, and corresponding low-level form. The articles such as the, an, a in the query are ignored. These words are then sorted according to priority and sent to the next level as given in Algorithm 3.

Table 1 Ipv6 address format

Base station Id	Level Id	Controller Id	Unused bits	Type Id	Value Id
8 bits	8 bits	8 bits	24 bits	8 bits	8 bits

Algorithm 2

```
Input : Query from user, q KTable-keywords and predefined set of
rules
Output: A prioritized sub query table.
     1: k ← Number of words in q, Table ← (i,j)
        where i is the word and j is the corresponding priority
     2: for each word w in q
     3:       if (isKeyword(w))
     4:            x ← assignPriority(w)
     5:            Table(i,j) ← (w,x)
     6: end for
     7: Sort(table(k,k)) //according to priorities
```

The location bit will indicate a unique binary representation for all the possible locations. The keyword bit is used to denote type of device whose value is requested by the user. The value bits will either indicate the specific device number (system no. 59) or the sensor value that the user wants to search (temperature 98 °C). It can be differentiated by using the flag bit. If flag bit is 1, it denotes the ID of the device else it denotes the specific value produced by the device. The S bit will denote whether the user wants an aggregated value or a single value. If it is 1, it denotes that aggregate value is needed.

Algorithm 3

```
Input : Prioritized subquery table, t
Output: Binary value of query
     1: binary [ ] ← array which will contain the
                binary value of the query
     2: For each word w in table
     3:       if priority (w) < 3
     4:            binary [] ← Search (w, ontology)
     5:       if priority (w) = 3
     6:            binary [ ] ← w
     7:            set next bit of binary as 1
     8:       if priority (w) = 4
     9:            binary [] ← w
     10:           set next bit of binary as 0
     11:      w ← next word in table t
     12:      if priority (w) =5
     13:         set next of binary as 1
     14: Send binary value to the base station.
```

Table 2 List of keywords

Type	Words	Priority	ID	Type	Words	Priority	ID
Location	Dcl lab	1		Id	5	3	0000000101
Location	Hpc lab	1	0011	Id	20	3	0000010100
Keyword	Temperature	2	0000	Value	45 °C	4	0000101101
Keyword	Light	2	0001	Value	5 V	4	0000000101

Table 3 Example query 1 and 2

Query	Dcl lab	Temperature	45 °C	Id/value	All
1	1	2	4	–	5
	0001	0000	0000101101	0	1
Query	Hpc lab	System	20	Id/value	s
2	1	2	3	–	–
	1000	0011	0000000011	1	0

For illustration purposes, let us consider an example query given by user. This query requests an aggregated data from the system. The list of keywords in the given environment is given in Table 2. Comparing given query with the predefined ontology, the query is transformed into a low-level binary form as shown in Table 3. Example query 2 requests a single value from heat sensor present near system no. 20.

Query 1: Find all temperature sensors which has 45 °C in Dcl lab and find whether the system no. 20 in Hpc lab is on.

3.5 Search

The search can be performed by comparing the bits in the transformed query with the device identifiers which we assigned earlier. It is a recursive process and formation of clusters will significantly reduce the search space. Using, the transformed query and the sender address, we will first go to the specific base station. Then, the query is compared with the controller id's present in the base station. If it matches, go to that specific controller. This process continues, until the required node is. We follow depth-first search here, as we reach the required node by following a relevant branch as given in Algorithm 4. If the user requests an aggregate value, we keep track of all the relevant data while searching and return the aggregated value.

Algorithm 4

```
Inputs: Set of sensors, Transformed query Value := 0
Output: The value requested by the user
            1: Go to the base station, BS
            2: For all controllers C within BS
            3:  If Controller ID == ID in query
            4:  If flag ==1
            5:   Value = Value + Controller data
            6:  If specific sensor is reached
            7:   Return Value
            8:  Else
            9:   Recursively search in DFS manner
            10: End For
```

4 Results and Discussion

The proposed system is implemented using Cooja simulator. Cooja is a network simulator operated by contiki os. The query processing is done using c programming language. And the processed query file is given as input to the Cooja simulator. The IoT environment is simulated by considering the sensors, such as temperature, light, heat, humidity, etc. we considered college department as scenario where all the necessary sensors are placed in each room/lab. 200 heat sensors, 100 temperature sensors, 100 humidity sensors, and 100 light sensors are considered for simulation. Table 4 shows the simulation parameters necessary to carry out implementation.

This section also analyzes the performance of the proposed system. All existing searching techniques search for a particular sensor in a sampling database, but our proposed system searches in real time. It is shown that from Fig. 2 response time for y search query with respect to Table 3 (query 1 and query 2). The proposed approach (PA) takes minimum response time because search take place based on BS ID, Level ID and Controller ID. But, generalized approach (GA) take long response time to search a node because less aware of identity of node position. Even increase in number of nodes in the networks, PA required more or less same response time for each keyword compare to GA. In GA each time increase the

Table 4 Simulation parameters

Network size (M × M)	200 m × 200 m
No. of sensors	500
Position of base station (x,y)	0,0
Communication range	15 m

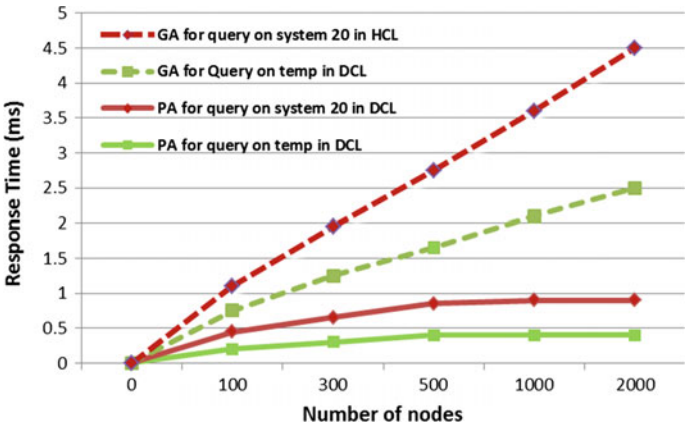


Fig. 2 Response time of sensor search

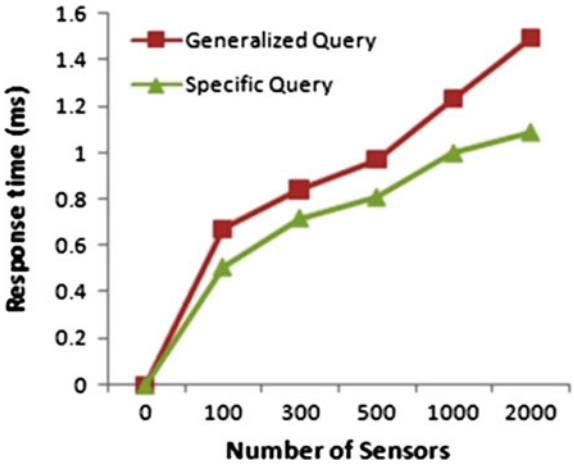


Fig. 3 Response time of the system with respect to different query types

response time of the node time increases gradually as number of sensors increases. This is because search space is greatly reduced by grouping the sensors in a cluster form.

The processing time required to convert the query to binary form is also less. The proposed system also handles queries of different type with respect to space and value. And the response time of the system not only depends on the number of sensors used but also on the type of query. The generalized queries may take longer time compared to specific queries this is because it has retrieve data from all sensors within that region. Figure 3 shows the performance of the system with respect to different queries. And, initially the startup time for the system will be higher because of cluster formation and addressing.

5 Conclusion

Since the sensor technology is growing and the costs are reduced, it is expected that every object around us will be attached to sensors. Thus, the search space of sensors will increase tremendously. In this paper, we defined a cluster-based search approach that will help us search the sensors efficiently by reducing the search space. We also allowed the users to enter queries in natural and abstract form which saved them from learning complex languages. In the future work, the clustering technique can be improved in a way that it supports Internet of Things.

References

1. Fredj, S.B., Boussard, M., Kofman, D., Noirie, L.: A scalable IoT service search based on clustering and aggregation. In: IEEE International Conference on Green Computing and Communications (2013)
2. Li, S., Da Xu, L., Zhao, S.: The internet of things: a survey. Springer Inf. Syst. Front. **17**, 243–259 (2015)
3. 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. 601–642 (2015)
4. Sarkar, C., Nambi, A.U., Venkatesha Prasad, R., Rahim, A., Neisse, R., Baldini, G.: DIAT: a scalable distributed architecture for IoT. IEEE J. **3** (2015)
5. Perera, C., Zaslavsky, A., Liu, C.H., Compton, M., Christen, P., Georgakopoulos, D.: Sensor search techniques for sensing as a service architecture for the Internet of Things. IEEE Sens. J. **14,** 406–419 (2014)
6. Elahi, B.M., Römer, K., Ostermaier, B.: Sensor ranking: a primitive for efficient content-based sensor search. In: ACM International Conference on Information Processing in Sensor Networks, pp. 217–228 (2009)
7. Truong, C., Romer, K., Chen, K.: Fuzzy based sensor search in the Web of things. In: IEEE International Conference on the Internet of Things (IOT) (2012)
8. Ding, Z, Gao, X., Guo, L., Yang, Q.: A hybrid search engine framework for the Internet of Things based on spatial-temporal, value-based, and keyword-based conditions. In: IEEE International Conference on Green Computing and Communications (2012)
9. Ebrahimi, M., Shafieibavani, E., Wong, R.K., Chi, C.H.: A new meta-heuristic approach for efficient search in the Internet of Things. In: IEEE International Conference on Services Computing (2015)
10. Yang, S., Xu, Y., He, Q.: Ontology based service discovery method for Internet of Things. In: IEEE International Conferences on Internet of Things, and Cyber, Physical and Social Computing (2011)

Assessment of Occupational Risks in Construction Sites Using Interval Type-2 Fuzzy Analytic Hierarchy Process

Joy Debnath and Animesh Biswas

Abstract This paper describes a method for assessing risk of the workers at construction sites using interval type-2 fuzzy analytic hierarchy process. Historical accident data, subjective judgments by the experts and relative importance of the risk factors are combined together to determine the current risk level of construction sites. The linguistic terms associated with the model are represented by interval type-2 trapezoidal fuzzy numbers. The proposed method can identify the risk factors which are most important in improving worker safety and, therefore, determines the areas on which the management should emphasize in order to improve the safety of the workers. The application potentiality of this model has been demonstrated by a case example of risk analysis of a construction industry.

Keywords Risk assessment · Linguistic variables · Type-2 trapezoidal fuzzy numbers · Interval type-2 fuzzy analytic hierarchy process

1 Introduction

Risk assessment is an essential and systematic process for assessing the impact and rate of hazards on human activities. It constitutes a needful tool for the safety policy of an industry. Pinto et al. [1] developed an occupational risk assessment technique in construction sites using traditional methods and fuzzy set theoretic approaches. Considering the expenses in health care, safety training, upgrading the machines and safety tools, a risk assessment method was developed by Beriha et al. [2] in the recent past. Furthermore, Thomas et al. [3] proposed a risk probability and impact assessment framework for risk assessment based on the fuzzy-fault tree and the Delphi method to deal with inadequate data.

J. Debnath · A. Biswas (✉)
Department of Mathematics, University of Kalyani, Kalyani 741235, India
e-mail: abiswaskln@rediffmail.com

J. Debnath
e-mail: joy6debnath@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_28

Due to the difficulty in adapting complex situation, a number of quantitative and qualitative risk analysis methods have been developed. In this circumstance, the analytic hierarchy process (AHP) [4] is appeared as one of the most widely used multi-criteria decision-making (MCDM) technique which can combine qualitative and quantitative factors for ranking and evaluating alternatives. Decision makers usually prefer to use linguistic variables rather than expressing their judgments in the form of exact numeric values. In such situations, fuzzy AHP (FAHP) is found as a useful tool to deal with imprecise, uncertain, or ambiguous data [5].

The concept of a type-2 fuzzy set (T2 FS) was introduced by Zadeh [6] as a generalization of the concept of an ordinary fuzzy set called type-1 fuzzy set (T1 FS). With the development of T2 FSs and with the initiation of the concept of interval type-2 fuzzy sets (IT2 FSs) [7], decisions received more flexibility in representing uncertainties. It is evident that the computational complexities associated with IT2 FSs are lesser [8] to a larger extent than that of T2 FSs. Thus, IT2 FSs became most acceptable fuzzy sets in representing uncertainties because of their relative simplicity than T2 FSs.

Abdullah and Najib [9] developed interval type-2 FAHP (IT2 FAHP) to solve various MCDM problems [10]. Wang et al. [11] investigated multi-attribute decision-making (MADM) problems in an IT2 fuzzy environment and developed an approach to handle the situation where the attribute values are partially known.

In the present study, almost all probable causes of accidents those are observed at job sites of the construction industries are taken into consideration and are identified as risk factors. To deal with the imprecise and inadequate historical data, IT2 FAHP is used to evaluate the weights of each risk factor. Again the risk scores for each factor that expresses the type of damages done previously are assigned by safety experts on the basis of their expertise. The overall risk level score value is then evaluated in order to assess the present work condition. Finally, the larger applicability of the extended agile dimensions of IT2 AHP than FAHP method is conveyed by comparing the associated results through a case study.

2 Some Basic Concepts

In this section, some basic concepts of IT2 FSs and arithmetic operations on trapezoidal IT2 FSs which are used in this chapter are discussed.

2.1 T2 FS

T2 FS is a fuzzy set whose membership functions are themselves T1 FSs. A T2 FS, $\tilde{\tilde{A}}$, in the universe of discourse, *X,* can be represented by a type-2 membership function $\mu_{\tilde{\tilde{A}}}(x)$ as shown below [12]:

$$\tilde{\tilde{A}} = \left\{ \left((x,u), \mu_{\tilde{\tilde{A}}}(x,u) \right) : x \in X, \quad u \in J_x \subseteq [0,1], \quad 0 \le \mu_{\tilde{\tilde{A}}}(x,u) \le 1 \right\} \tag{1}$$

where J_x denotes a subinterval in [0, 1].

A T2 FS, $\tilde{\tilde{A}}$, can also be represented as follows:

$$\tilde{\tilde{A}} = \int_{x \in X} \int_{u \in J_x} \mu_{\tilde{\tilde{A}}}(x,u)/(x,u) \quad \text{where } J_x \subseteq [0,1]. \tag{2}$$

2.2 IT2 FS

Let $\tilde{\tilde{A}}$ be a T2 FS in the universe of discourse, *X,* represented by the type-2 membership function $\mu_{\tilde{\tilde{A}}}$. If $\mu_{\tilde{\tilde{A}}}(x,u) = 1$ for all points (x,u), then $\tilde{\tilde{A}}$ is called IT2 FS. It is represented as follows [8]:

$$\tilde{\tilde{A}} = \int_{x \in X} \int_{u \in J_x} 1/(x,u) \quad \text{where } J_x \subseteq [0,1]. \tag{3}$$

2.3 Trapezoidal IT2 FS

The upper and lower membership functions of an IT2 FS are type-1 FSs. Let $\tilde{\tilde{A}}$ be a trapezoidal IT2 FS (Fig. 1) defined on the universe of discourse *X*. Then $\tilde{\tilde{A}}$ can be illustrated as follows:

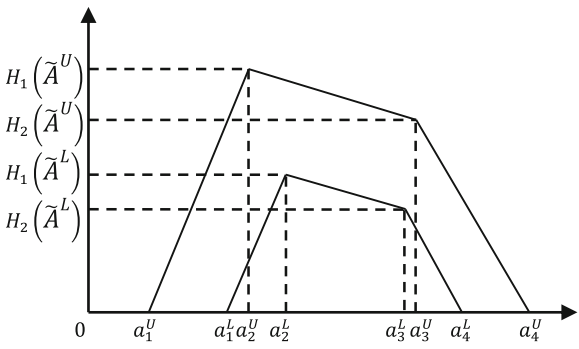


Fig. 1 Trapezoidal IT2 FS

$$\tilde{\tilde{A}} = \left(\tilde{A}^{U}, \tilde{A}^{L}\right) = \begin{pmatrix} \left(a_1^U, a_2^U, a_3^U, a_4^U; H_1\left(\tilde{A}^{U}\right), H_2\left(\tilde{A}^{U}\right)\right), \\ \left(a_1^L, a_2^L, a_3^L, a_4^L; H_1\left(\tilde{A}^{L}\right), H_2\left(\tilde{A}^{L}\right)\right) \end{pmatrix}$$

where, $H_1\left(\tilde{A}^{U}\right)$ and $H_2\left(\tilde{A}^{U}\right)$ denote the respective membership value at the points a_2^U and a_3^U in the upper trapezoidal membership function $\tilde{A}^{U}$. $H_1\left(\tilde{A}^{L}\right)$ and $H_2\left(\tilde{A}^{L}\right)$ denote the membership grades of the points a_2^L and a_3^L, respectively, in the lower trapezoidal membership function $\tilde{A}^{L}$.

2.4 Arithmetic Operations on Trapezoidal IT2 FS

Let

$$\tilde{\tilde{A}}_1 = \begin{pmatrix} \left(a_{11}^U, a_{12}^U, a_{13}^U, a_{14}^U; H_1\left(\tilde{A}_1^U\right), H_2\left(\tilde{A}_1^U\right)\right), \\ \left(a_{11}^L, a_{12}^L, a_{13}^L, a_{14}^L; H_1\left(\tilde{A}_1^L\right), H_2\left(\tilde{A}_1^L\right)\right) \end{pmatrix}, \tilde{\tilde{A}}_2 = \begin{pmatrix} \left(a_{21}^U, a_{22}^U, a_{23}^U, a_{24}^U; H_1\left(\tilde{A}_2^U\right), H_2\left(\tilde{A}_2^U\right)\right) \\ \left(a_{21}^L, a_{22}^L, a_{23}^L, a_{24}^L; H_1\left(\tilde{A}_2^L\right), H_2\left(\tilde{A}_2^L\right)\right) \end{pmatrix}$$

be two IT2 FSs.

The arithmetic operations of trapezoidal IT2 FSs [13, 14] are illustrated as follows.

2.4.1 Addition

The addition of two trapezoidal IT2 FSs is denoted by, $\tilde{\tilde{A}}_1 \oplus \tilde{\tilde{A}}_2$, and is defined as

$$\tilde{\tilde{A}}_1 \oplus \tilde{\tilde{A}}_2 = \begin{pmatrix} \left(a_{11}^U + a_{21}^U, a_{12}^U + a_{22}^U, a_{13}^U + a_{23}^U, a_{14}^U + a_{24}^U, \min\left(H_1\left(\tilde{A}_1^U\right), H_1\left(\tilde{A}_2^U\right)\right), min\left(H_2\left(\tilde{A}_1^U\right), H_2\left(\tilde{A}_2^U\right)\right)\right), \\ \left(a_{11}^L + a_{21}^L, a_{12}^L + a_{22}^L, a_{13}^L + a_{23}^L, a_{14}^L + a_{24}^L, \min\left(H_1\left(\tilde{A}_1^L\right), H_1\left(\tilde{A}_2^L\right)\right), \min\left(H_2\left(\tilde{A}_1^L\right), H_2\left(\tilde{A}_2^L\right)\right)\right) \end{pmatrix} \tag{4}$$

2.4.2 Multiplication

The multiplication between the two trapezoidal IT2 FS is denoted by, $\tilde{\tilde{A}}_1 \otimes \tilde{\tilde{A}}_2$, and is defined as

$$\tilde{\tilde{A}}_1 \otimes \tilde{\tilde{A}}_2 = \begin{pmatrix} \left(a_{11}^U \times a_{21}^U, a_{12}^U \times a_{22}^U, a_{13}^U \times a_{23}^U, a_{14}^U \times a_{24}^U, \min\left(H_1\left(\tilde{A}_1^U\right), H_1\left(\tilde{A}_2^U\right)\right), min\left(H_2\left(\tilde{A}_1^U\right), H_2\left(\tilde{A}_2^U\right)\right)\right), \\ \left(a_{11}^L \times a_{21}^L, a_{12}^L \times a_{22}^L, a_{13}^L \times a_{23}^L, a_{14}^L \times a_{24}^L, \min\left(H_1\left(\tilde{A}_1^L\right), H_1\left(\tilde{A}_2^L\right)\right), min\left(H_2\left(\tilde{A}_1^L\right), H_2\left(\tilde{A}_2^L\right)\right)\right) \end{pmatrix} \tag{5}$$

3 Proposed Methodology

In comparison with the variants of AHP [4], IT2 FAHP is seemed as a more systematic and proficient tool in order to deal with uncertain data. Also, IT2 FAHP has its greater capability of capturing data in a more efficient manner from any linguistic hedges than other methods. In the proposed process of risk assessment using IT2 FAHP, first of all, a hierarchy is prepared by considering all possible risk factors in the work environment. By evaluating the weights of different kinds of risk factors, the overall risk level score of the industry is evaluated. Finally, a suitable safety development policy for the workers is discussed. The proposed methodology is presented below.

3.1 Knowledge Acquisition and Data Collection

The collection and analysis of data helps in better understanding of different types of accidents occurred on different construction sites. The records in the archive and the expert witness are important resources for the data collection, which includes preliminary official records, statements made by the injured workers and his/her relatives, statements made by site engineers, safety managers, etc.

3.2 Identification of Risk Factors

On the basis of the collected data, all potential risk factors associated with a construction site are identified. Based on the literature, historical data and past experiences, the risk factors in the construction sites are listed to assess the risk in construction sites for the proposed methodology.

3.3 Hierarchical Model for Risk Assessment

On the basis of the collected data, all possible causes of various kinds of injuries of workers on the job site are identified and then those are presented in a form of a hierarchy.

3.4 Evaluation of Weight of Risk Factors

The possible influence of a risk factor to any accidents is measured by calculating its relative weight. The following method is proposed to determine the weight of each factor.

3.4.1 Methodology for Weight Evaluation Using IT2 FAHP

AHP method is further generalized here by considering the inputs in the form of interval type 2 fuzzy numbers (IT2 FN) for better reflection of uncertainties. The whole process of this method is explained through the following steps.

Step 1 The pairwise comparisons of the criteria are made using IT2 trapezoidal fuzzy comparison scale which is described in Table 1. The results of the comparisons are described as fuzzy pairwise comparison matrices as follows:

$$\tilde{\tilde{C}} = \left[\tilde{\tilde{C}}_{ij}\right] = \begin{bmatrix} \tilde{\tilde{C}}_{11} & \tilde{\tilde{C}}_{12} & \cdots & \tilde{\tilde{C}}_{1n} \\ \tilde{\tilde{C}}_{21} & \tilde{\tilde{C}}_{22} & \cdots & \tilde{\tilde{C}}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\tilde{C}}_{n1} & \tilde{\tilde{C}}_{n2} & \ldots & \tilde{\tilde{C}}_{nn} \end{bmatrix}, \quad i,j = 1,2,\cdots,n \tag{6}$$

where $\tilde{\tilde{C}}_{ij} = \begin{pmatrix} \left(p_{ij}^U, q_{ij}^U, r_{ij}^U, s_{ij}^U; H_1\left(q_{ij}^U\right), H_2\left(r_{ij}^U\right)\right), \\ \left(p_{kl}^L, q_{kl}^L, r_{kl}^L, s_{kl}^L; H_1\left(q_{kl}^L\right), H_2\left(r_{kl}^L\right)\right) \end{pmatrix}$ represents the quantified judgment on c_i to c_j;

Table 1 Scale of relative importance for pairwise comparison

Trapezoidal fuzzy number	Linguistic variable
(1, 1, 1, 1; 1, 1) (1, 1, 1, 1; 1, 1)	Just equal (JE)
(1, 2, 3, 4; 1, 1) (1.4,2.4, 2.6, 3.6; 0.8, 0.8)	Weakly important (WI)
(2, 3, 4, 5; 1, 1) (2.4, 3.4, 3.6, 4.6; 0.8, 0.8)	Between weakly and strongly important (BWSI)
(3, 4, 5, 6; 1, 1) (3.4, 4.4, 4.6, 5.6; 0.8, 0.8)	Strongly important (SI)
(4, 5, 6, 7; 1, 1) (4.4, 5.4, 5.6, 6.6; 0.8, 0.8)	Between strongly and very strongly important (BSVI)
(5, 6, 7, 8; 1, 1) (5.4, 6.4, 6.6, 7.6; 0.8, 0.8)	Very strongly important (VSI)
(6, 7, 8, 9; 1, 1) (6.4, 7.4, 7.6, 8.6; 0.8, 0.8)	Between very strongly and absolutely important (BVAI)
(7, 8, 9, 9; 1, 1) (7.4, 8.4, 8.6, 9; 0.8, 0.8)	Absolutely important (AI)

$\tilde{\tilde{C}}_{ji} = \begin{pmatrix} \left(\frac{1}{s_{ij}^U}, \frac{1}{r_{ij}^U}, \frac{1}{q_{ij}^U}, \frac{1}{p_{ij}^U}; H_1\left(q_{ij}^U\right), H_2\left(r_{ij}^U\right)\right), \\ \left(\frac{1}{s_{ij}^L}, \frac{1}{r_{ij}^L}, \frac{1}{q_{ij}^L}, \frac{1}{p_{ij}^L}; H_1\left(q_{kl}^L\right), H_2\left(r_{kl}^L\right)\right) \end{pmatrix}$ represents the quantified judgment on c_j to c_i and n represents the number of criteria.

Step 2

Before calculating the weights, the consistency of the comparison matrix has to be checked. In order to check the consistency of the comparison matrix, intuitively, the IT2 FNs are to be converted first into its equivalent defuzzified values.

The defuzzification of a trapezoidal IT2 fuzzy (DTraT) set as described in Fig. 1 is evaluated by Eq. (7) [15].

$$\text{DTraT} = \frac{\frac{\left(a_4^U - a_1^U\right) + \left(H_2\left(\tilde{A}^U\right) \cdot a_2^U - a_1^U\right) + \left(H_1\left(\tilde{A}^U\right) \cdot a_3^U - a_1^U\right)}{4} + l_U + \frac{\left(a_4^L - a_1^L\right) + \left(H_2\left(\tilde{A}^L\right) \cdot a_2^L - a_1^L\right) + \left(H_1\left(\tilde{A}^L\right) \cdot a_3^L - a_1^L\right)}{4} + l_L}{2} \tag{7}$$

Based on the defuzzification as discussed above, the fuzzy comparison matrix $\tilde{\tilde{C}}$ is converted into the crisp comparison matrix. The consistency of the crisp comparison matrix is calculated by the usual method. It is to be noted here that if the consistency ratio of that matrix is less than 0.10, the consistency of the matrix $\tilde{\tilde{C}}$ is considered as acceptable, otherwise the pairwise comparisons should be revised.

Step 3 The fuzzy weights are calculated using geometric mean method [16]. The geometric mean $\tilde{\tilde{C}}_i$ of the ith row in the comparison matrix is defined as

$$\tilde{\tilde{C}}_i = [\tilde{\tilde{C}}_{i1} \otimes \cdots \otimes \tilde{\tilde{C}}_{in}]^{1/n} \tag{8}$$

where

$$\sqrt[n]{\tilde{\tilde{C}}_{ij}} = \begin{pmatrix} \left(\sqrt[n]{p_{ij}^U}, \sqrt[n]{q_{ij}^U}, \sqrt[n]{r_{ij}^U}, \sqrt[n]{s_{ij}^U}; H_1\left(q_{ij}^U\right), H_2\left(r_{ij}^U\right)\right), \\ \left(\sqrt[n]{p_{kl}^L}, \sqrt[n]{q_{kl}^L}, \sqrt[n]{r_{kl}^L}, \sqrt[n]{s_{kl}^L}; H_1\left(q_{kl}^L\right), H_2\left(r_{kl}^L\right)\right) \end{pmatrix} \tag{9}$$

Step 4 The fuzzy weight of the ith criteria is evaluated as

$$w_i = \tilde{\tilde{C}}_i \otimes [\tilde{\tilde{C}}_1 \oplus \cdots \oplus \tilde{\tilde{C}}_i \oplus \cdots \oplus \tilde{\tilde{C}}_n]^{-1} \tag{10}$$

Step 5 Fuzzy weights are converted into crisp weights (w_i^*) by using defuzzification method as defined in Eq. (7).

Step 6 Crisp weights are normalized using the following equation as

where w_i is the weight of c_i after normalization process.

$$w_i = \frac{w_i^*}{\sum_{i=1}^{n} w_i^*}, \tag{11}$$

3.5 Assessment of Overall Risk Level Score

The overall risk level score (RLS) of a construction industry is calculated by the following equation.

$$RLS_{OVERALL} = \sum_{i}^{n} \left(W_i . \sum_{j}^{k_i} w_{ij} r_{ij} \right) \tag{12}$$

where r_{ij} is the risk score (RS) given for each risk factor in second level of the hierarchy which expresses the extent of damages incurred due to those factors. W_i and w_{ij} are the respective weights of the associated risk factors in the first and second level.

The developed process is implemented on a construction industry to assess the overall risk and thereby identifying those risk factors which are predominantly responsible for the accidents in the following section.

4 Assessment of Risk on a Construction Site: A Case Study

An Indian construction industry is chosen to demonstrate the proposed methodology. On the basis of the collected data, all the causes of accidents occurred during the last 20 years is classified and presented through the following hierarchical diagram.

In the present case study, the database is consisting of all the causes of different kind of accidents took place on job sites in that industry as presented in Fig. 2. Considering these causes, a hierarchy of overall RSL is constructed on the basis of experts' knowledge. In the first level of the hierarchy seven risk factors are identified, such as *unsafe handling of material (R_1), poor housekeeping (R_2), lack of communication (R_3), flying particles (R_4), poor supervision (R_5), lack of knowledge (R_6), human factors (R_7).* At the second level all the risk factors in the first level are divided into subfactors.

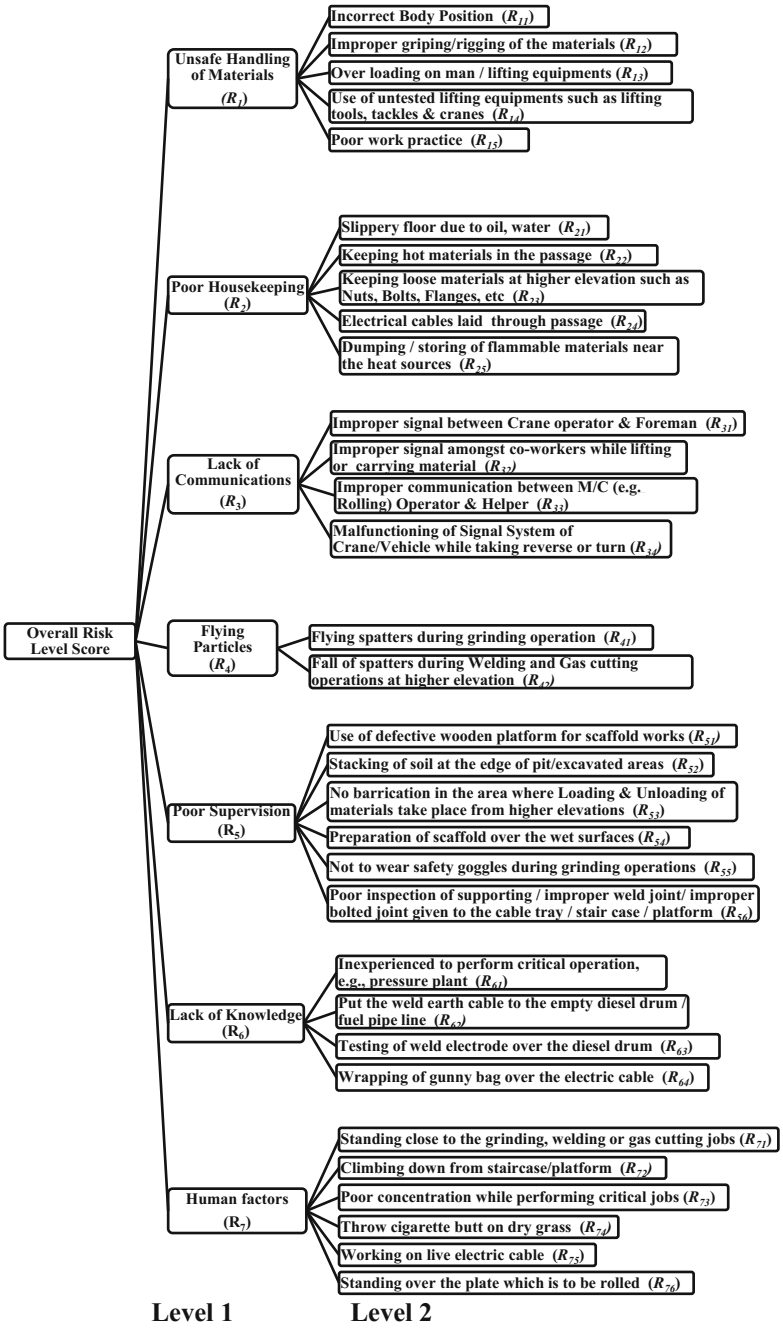


Fig. 2 Hierarchical structure of RLS

On the basis of experts' expertise and job site inspections, *unsafe handling of materials (R_1)* is classified into five subfactors which are Incorrect body position (R_{11}), improper griping/rigging of the materials (R_{12}), overloading on man/lifting equipment (R_{13}), use of untested lifting equipment, such as lifting tools and tackles and cranes (R_{14}), poor work practice (R_{15}). It is observed that the workers are very frequently injured for these factors on foot, fingers, chest, and eyes.

Proper housekeeping plays an important role in workers' safety on job sites. As a consequence, slippery floor due to oil, water (R_{21}), keeping hot materials in the passage (R_{22}), keeping loose materials at higher elevation, such as nuts, bolts, flanges, etc. (R_{23}), electrical cables lay through passages (R_{24}), and dumping/storing of flammable materials near the heat sources (R_{25}) causes various injuries to lower leg, head, foot, etc. In the present study, these causes are considered as *poor housekeeping (R_2)*.

It is seen that, the modern mechanism cannot always control hazards which are being occurred due to *Lack of communications (R_3)* between man–machine interface. This risk factor occurs for Improper signal between crane operator and foreman (R_{31}), improper signal amongst coworkers while lifting or carrying material (R_{32}), improper communication between M/C (e.g., Rolling) operator and helper (R_{33}), and malfunctioning of signal system of crane/vehicle while taking reverse or turn (R_{34}). The most injury prone body parts corresponding to this risk factor are identified as head, hands, and fingers.

According to the historical data, it is evidenced that *flying particles (R_4)* like flying spatters during grinding operation (R_{41}) and fall of spatters during welding and gas cutting operations at higher elevation (R_{42}) causes injuries on thigh, lower leg, and forearm to the workers of the industry.

Now a day's *poor supervision (R_5)* in construction sites becomes an inevitable risk factor. In this model, Use of defective wooden platform for scaffold works (R_{51}), stacking of soil at the edge of pit/excavated areas (R_{52}), no barricading in the area where loadingand unloading of materials take place from higher elevations (R_{53}), preparation of scaffold over the wet surfaces (R_{54}), not to wear safety goggles during grinding operations (R_{55}), Poor inspection of supporting/improper weld joint/improperly bolted joint given to the cable tray/stair case/platform (R_{56}) are characterized as the subfactors of R_5.

Modern machineries and technology sometimes cannot protect workers from several injuries occurred due to *Lack of knowledge (R_6)* in the form of Inexperienced to perform critical operation, e.g., pressure plant (R_{61}), Ppt the weld earth cable to the empty diesel drum/fuel pipeline (R_{62}), testing of weld electrode over the diesel drum (R_{63}), and wrapping of gunny bag over the electric cable (R_{64}).

Strict safety protocol and upgraded safety tools may not always reduce various severe injuries occurred to the workers due to *Human factors* (R_7), viz., Standing close to the grinding, welding, or gas cutting jobs (R_{71}), climbing down from staircase/platform (R_{72}), poor concentration while performing critical jobs like bending work in steel yard, passing through the work area, scaffold alignment, wood cutting, etc. (R_{73}), throwing cigarette butts on dry grass (R_{74}), Working on live electric cable (R_{75}), standing over the plate which is to be rolled (R_{76}).

As a result, every year a large number of workers suffer mostly from eye, head, lower leg, foot, finger, forearm, and pelvis injuries.

In the evaluation process of the weights of each risk factor and associated subfactors described in Fig. 2, the initial pairwise comparison matrices are constructed using experts' opinion based on the preference scale as shown in 1. The weights of each risk factors and subfactors are then calculated by executing each step as described in the Sect. 3.4.1. The initial pairwise comparison matrix of all risk factors in the level 1is presented in the following Table 2.

4.1 Results and Discussions

The evaluation of overall RLS of the industry has been performed using the process as described in Sect. 3.5 and the achieved results are then compared with the output of FAHP approach. The consequences of weights of all the factors in Level 1 achieved in the form of IT2 FNs are shown in Table 3.

Similarly, the evaluation of the weights of each subfactors associated with each risk factor in the form of IT2 FNs has been executed. The calculated IT2 FNs are then defuzzified using the defuzzification method as described in Step-5.

Table 2 Initial comparison matrix of risk factors in the level-1

Risk factors → ↓	R_1	R_2	R_3	R_4	R_5	R_6	R_7
R_1	*JI*	*SI*	*BVAI*	*AI*	*BSVI*	*VSI*	1/WI
R_2	1/*SI*	*JI*	*VSI*	*BVAI*	*SI*	*BWSI*	1/*SI*
R_3	1/*BVAI*	1/*VSI*	*JI*	*WI*	1/*SI*	1/*BWAI*	1/*BVAI*
R_4	1/*AI*	1/*BVAI*	1/*WI*	*JI*	1/*BSVI*	1/*BWSI*	1/*AI*
R_5	1/*BSVI*	1/*SI*	*SI*	*BSVI*	*JI*	*BWSI*	1/*BSVI*
R_6	1/*VSI*	1/*BWSI*	*BWSI*	*BWSI*	1/*BWSI*	*JI*	1/*VSI*
R_7	*WI*	*SI*	*BVAI*	*AI*	*BSVI*	*VSI*	*JI*

Table 3 Weights of the risk factors in the form of IT2 FNs

Risk factor	Type-2 fuzzy weight
R_1	(0.17, 0.24, 0.35, 0.51; 1, 1) (0.20, 0.28, 0.30, 0.43; 0.8, 0.8)
R_2	(0.09, 0.12, 0.18, 0.26; 1, 1) (0.10, 0.14, 0.15, 0.22; 0.8, 0.8)
R_3	(0.02, 0.02, 0.03, 0.05; 1, 1) (0.02, 0.03, 0.03, 0.04; 0.8, 0.8)
R_4	(0.01, 0.02, 0.02, 0.04; 1, 1) (0.01, 0.02, 0.02, 0.03; 0.8, 0.8)
R_5	(0.05, 0.07, 0.10, 0.15; 1, 1) (0.06, 0.08, 0.09, 0.13; 0.8, 0.8)
R_6	(0.03, 0.04, 0.06, 0.10; 1, 1) (0.03, 0.05, 0.05, 0.08; 0.8, 0.8)
R_7	(0.21, 0.32, 0.45, 0.63; 1, 1) (0.25, 0.36, 0.39, 0.54; 0.8, 0.8)

After evaluating the defuzzified weights of all risk factors and subfactors, the RS corresponding to each risk subfactors in the Level 2 is calculated. To calculate the RS, experts' judgments are taken into consideration on the basis of preference scale as shown in Table 4. Then, the overall RLS of the concerned industry is evaluated using (12). The quantitative results of the weights, RS of the risk factors and the overall RLS of the industry are consolidated and presented in Table 5.

The overall RLS of the concerned industry using the proposed methodology is obtained as 7.86, which indeed conveys the present risk level of that industry and is also very much comprehensible for the administration to locate the current status.

In order to compare the proposed method with FAHP method, the risk assessment process is re-executed using type-1 fuzzy numbers with the same data that has been used in IT2 FAHP method and is presented in Table 5 with the results of the developed IT2 FAHP.

From Table 5 it is observed that according to FAHP the relative weights of R_{11} and R_{14}; R_{24} and R_{24}; R_{31} and R_{34}; R_{71}, R_{72}; and R_{74} are equal. But executing the proposed model, these inconsistencies are reduced by providing different score values.

Again this case study clearly indicates the significance of using IT2 FAHP by considering the small comparable importance between the factors. It is also very much relevant to the safety personals and the administration of the industries for identifying the risk factors according to their priority level and to maintain efficiently the safety management processes. Further, it is clear from both the approaches that the risk factors R_7, R_1 and R_2 are predominantly responsible for the accidents of workers in the construction sites. Therefore, it is advisable to the concerned industry that the occupational risks can be reduced by implementing efficient preventive measures and upgrading safety equipment related to those risk factors as mentioned above along with their associated subfactors.

Table 4 Preference scale for risk score

Scale	Risk score
First aid case/No loss time incident case (LTI)/No medical treatment case (MTC)	1
Case referred to Medical Centre/Hospital—No LTI/No MTC	2–3
Case referred to Medical Centre/Hospital, leading to either MTC or LTI	4–5
Partial permanent disablement	6
Total permanent disablement	7
Multiple injuries	8
Fatality	9
Multiple fatality	10

Table 5 Quantitative results of the weights, RS, and Overall RLS using FAHP and the Proposed IT2 FAHP

Risk factor	W_i		Subfactor	w_{ij}		RS (r_{ij})	Overall RLS	
	FAHP	IT2 FAHP		FAHP	IT2 FAHP		FAHP	IT2 FAHP
R_1	0.30	0.29	R_{11}	0.04	0.03	5	8.12	7.86
			R_{12}	0.37	0.36	7		
			R_{13}	0.31	0.32	8		
			R_{14}	0.04	0.05	6		
			R_{15}	0.24	0.24	9		
R_2	0.14	0.15	R_{21}	0.11	0.32	9		
			R_{22}	0.09	0.06	7		
			R_{23}	0.60	0.42	8		
			R_{24}	0.10	0.08	7		
			R_{25}	0.10	0.11	7		
R_3	0.03	0.03	R_{31}	0.13	0.12	7		
			R_{32}	0.70	0.67	9		
			R_{33}	0.04	0.05	4		
			R_{34}	0.13	0.16	6		
R_4	0.02	0.02	R_{41}	0.10	0.11	4		
			R_{42}	0.90	0.90	9		
R_5	0.08	0.08	R_{51}	0.03	0.03	4		
			R_{52}	0.49	0.47	9		
			R_{53}	0.10	0.09	6		
			R_{54}	0.05	0.05	5		
			R_{55}	0.09	0.12	6		
			R_{56}	0.24	0.24	8		
R_6	0.04	0.05	R_{61}	0.74	0.70	9		
			R_{62}	0.10	0.11	5		
			R_{63}	0.08	0.09	4		
			R_{64}	0.09	0.10	3		
R_7	0.39	0.37	R_{71}	0.06	0.06	7		
			R_{72}	0.06	0.08	8		
			R_{73}	0.57	0.53	9		
			R_{74}	0.06	0.06	6		
			R_{75}	0.16	0.19	8		
			R_{76}	0.09	0.08	6		

5 Conclusions

In this paper, the uncertainties in the risk assessment process have been incorporated with greater extent by using IT2 FAHP approach. This method provides generously proportioned flexibility than previous FAHP techniques which is

reflected by comparing the achieved results. This approach provides a new scientific way for risk assessment process which makes results more reasonable and comprehensive.

The developed model for evaluating industrial risks emphasizes on all possible hazardous causes of accidents. Based on the achieved results, this model shows its potential advantage in detecting most harmful risk factors systematically and effectively. It is evident that the proposed model not only reduces manpower investment but also mitigate the risk associated with workers on job sites of construction industries. Further, IT2 FSs, due to its third dimension in its MFs, provides more degrees of freedom for handling uncertainties. Such sets are useful in these circumstances, where it is difficult to determine the exact MFs of the linguistic variables.

Again, the proposed IT2 FAHP method can further be generalized by type-2 fuzzy algorithm in the learning process of input values. However, it is hoped that the developed technique may open up new vistas in the way of assessing occupational risks in construction sites.

Acknowledgements The authors are thankful to the anonymous reviewers for their comments and suggestions to improve the quality of the paper. The authors express their humble gratitude to Mr. Krishna Nirmalya Sen, President, American Society of Safety Engineers, India Chapter, for his expert opinion and kind cooperation in the process of execution of the developed model. This work is partially supported by DST-PURSE Programme of University of Kalyani, Kalyani, India.

References

1. Pinto, A., Nunes, I.L., Ribeiro, R.A.: Occupational risk assessment in construction industry—overview and reflection. Saf. Sci. **49**, 616–624 (2011)
2. Beriha, G.S., Patnaik, B., Mahapatra, S.S., Padhee, S.: Assessment of safety performances in Indian industries using fuzzy approach. Experts Syst. Appl. **39**, 3311–3323 (2012)
3. Thomas, A.V., Kalidindi, S.N., Ganesh, L.S.: Modelling and assessment of critical risks in BOT road projects. Constr. Manag. Econ. **24**, 407–424 (2006)
4. Saaty, T.L., Vargas, L.G.: The Analytic Hierarchy Process. McGraw-Hill, New York (1980)
5. Majumder, D., Debnath, J., Biswas, A.: Risk analysis in construction sites using fuzzy reasoning and fuzzy analytic hierarchy process. Proc. Technol. **10**, 604–614 (2013)
6. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)
7. Mendel, J.M.: Uncertain Rule-Based fuzzy Logic Systems: Introduction and New Directions. Prentice-Hall, Upper Saddle River, NJ (2001)
8. Mendel, J.M., John, R.I., Liu, F.L.: Interval type-2 fuzzy logical systems made simple. IEEE Trans. Fuzzy Syst. **14**, 808–821 (2006)
9. Abdullah, L., Najib, L.: A new type-2 fuzzy set of linguistic variables for the fuzzy analytic hierarchy process. Experts Syst. Appl. **41**, 3297–3305 (2014)
10. Mou, Q.: Method of Multi-Attribute Decision-Making and Its Application. Guangxi University, Nanning (2004)
11. Wang, W., Liu, X., Qin, Y.: Multi-attribute group decision making models under interval type-2 fuzzy environment. Knowl.-Based Syst. **30**, 121–128 (2012)
12. Jhon, R.I.: An appraisal of theory and applications. Int. J. Uncertain. Fuzziness Knowl. Based Syst. **6**, 563–576 (1998)

13. Lee, L.W., Chen, S.M.: A new method for fuzzy multiple attributes group decision-making based on the arithmetic operations of interval type-2 fuzzy sets. In: Proceedings of the 2008 International Conference on Machine Learning and Cybernetics, China, Kunming, pp. 3084–3089 (2008)
14. Chen, S.M., Lee, L.W.: Fuzzy multiple attributes group decision-making based on the interval type-2 TOPSIS method. Experts Syst. Appl. **37**, 2790–2798 (2010)
15. Kahraman, C., Oztaysi, B., Sari, I.U., Turanoglu, E.: Fuzzy analytic hierarchy process with interval type-2 fuzzy sets. Knowl. Based Syst. **59**, 48–57 (2014)
16. Ross, T.J.: Fuzzy Logic with Engineering Applications. McGraw-Hill, New York (1995)

Comparative Evaluation of Capacity Analysis and Probability Distribution of Elements for Different Iterative Values of MIMO

Sutanu Ghosh

Abstract In present scenario, MIMO is an attractive technology to enhance the system capacity using its antenna configuration. This capacity analysis is possible through the probability distribution of elements used in MIMO. In this work, a comparative analysis of probability distribution of elements has been done in a scattered environment of MIMO system. Different iterative values have been considered to perform this comparison. This time-based iterative value is useful to achieve the probability distribution with respect to different elements. Altogether, capacity comparison is also done through different antenna patterns. From the graphical output it is shown that, better level of distribution occurs at higher levels of iterative value.

Keywords MIMO correlation · SVD · Kronecker model · Weichselberger model · VCR model

1 Introduction

Recently, multi-antenna array technology is moreover popular due to improvement of system performance in terms of channel capacity enhancement, improvement of system quality of service and bandwidth utilization, error reduction, etc. Capacity of a channel is most significant to enhance the system throughput. Different antenna configurations in MIMO [1] are employed to modify the capacity linearly as per the system requirement. Conventional theoretical analysis has been done in sufficiently high scattering environments [2] to enhance the capacity gain using spatial beamforming or multiplexing in MIMO over SISO. Specially, the Ergodic capacity and spectral efficiency of the system is possible to enhance using equal amount of transmit

S. Ghosh (✉)
Indian Institute of Engineering Science and Technology, Shibpur, Howrah, India
e-mail: sutanu99@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_29

power among all transmit antennas [3]. It is also established [2–3] that Shannon capacity of an array, which containing n_t transmitting elements and an n_r receiving elements is roughly proportional to $\min(n_t, n_r)$ for large numbers of antennas. On the other hand, mutual information of MIMO is studied for spatially correlated channels and correlated noise in [4].

Multi-element antenna arrays (MEA) configuration at both sides of transmitter and receiver can support the spatial dimension to achieve higher channel capacity than traditional techniques. However, the architecture of transmission and reception of MEA configuration is much more complex than traditional antenna array. Accordingly, tractable capacity analysis is involved in realistic practical environment. Conversely, various physical attributes like spatial structure and correlation between transmitting and receiving antennas, channel complexity, interference from external sources, etc. have a great influence to modify the capacity of MIMO. One of the most important attributes is channel correlation to achieve more channel capacity through large number of correlated or decorrelated subchannels. Channel correlation is an evaluation of likeliness between two or, more different channels. Correlation can be expressed in terms of correlation factor, which is varying from 0 to 1. Higher level of correlation provides less number of subchannels, whereas large numbers of multiple channels are available for low level correlation. However, higher capacity can be achieved through large number of subchannels.

In this paper, a comparative capacity analysis of different MIMO configuration and probability distribution function in Singular Value Decomposition (SVD) is produced for different iterative values. Moreover, the high diversity gain in [5] is achieved through uniform channel decomposition. On the basis of channel condition, SVD-based channel decomposition has different outcomes due to different subchannels for a higher order of signal-to-noise ratio (SNR). To the best of my knowledge, before this work there is no such comparative study yet presented in this issue.

The remaining portions of this paper are arranged as follows: Sect. 2 describes MIMO and usage of SVD-based channel decomposition in MIMO. Section 3 gives the idea about MIMO channel correlation and the characteristics of different channel model in MIMO. Simulation and experimental results are illustrated in Sect. 4. Finally, this paper is concluded in Sect. 5.

2 MIMO Technology and SVD-Based Channel Decomposition

MIMO includes multiple antennas in transmitter and receiver with different levels of channel correlation. It is an interesting technology due to it's two important techniques—one is space time coding method by developing diversity gain [6–8] for communication reliability and the other is spatial multiplexing method [9] for

maximization of system throughput. It is to be assumed that receiver has proper channel state information (CSI) for both of these two methods. One of the most important methods is SVD, which can decompose a MIMO channel into a number of parallel subchannels. In this case, water filling algorithm is used to achieve high channel capacity [3] for the system. To provide the basic scheme of Generalized SVD (GVSD), first of all it is required to introduce the deterministic Gaussian channel model.

$$\begin{aligned} Y_r &= h_r X(t) + U_r(t), \\ Y_e &= h_e X(t) + U_e(t), \quad \text{for t} = 0, 1, 2, \ldots, \text{M} \end{aligned} \tag{1}$$

where, $h_r \in \mathbb{C}^{n_r \times n_t}$ and $h_e \in \mathbb{C}^{n_e \times n_t}$ are deterministic channel matrices for intended receiver and eavesdroppers. n_t, n_r and n_e are the number of transmitting antenna, intended receiving antenna and eavesdropper antenna, respectively. Noise vectors are represented by $U_r(t)$ and $U_e(t)$ for intended receiver and eavesdroppers, respectively. SVD is the special case of GSVD, which is not unitary; whereas SVD is fully applicable for unitary matrices. SVD of the system matrix M with rank m can be written as $M = U\sum V^H$, where U and V are the unitary matrices [10].

Importance of SVD in MIMO communication systems relies on the fact that the presteering matrix at transmitter and steering matrix at receiver decompose into m number of SISO orthogonal modes of excitation.

$$U^H Y = U^H MVX + U^H N = \sum X + U^H N \tag{2}$$

3 MIMO Channel Correlation and Different Channel Model of MIMO

Correlation is an important technique for MIMO technology to improve the channel capacity. Multiple antennas are to be utilized to differentiate the signal strength in MIMO. Multiple antenna array like 2 × 2, 4 × 4, etc. are used to fulfill the data transmission over the wireless path. Actually this 4 × 4 means both of the two sides of transmission have 4 numbers of antennas. These antenna array patterns have definite spatial correlation between transmitter and receiver end. Spatial correlation is purely a cross correlation of those signals that initiate from same source and is found at two spatially separated antenna elements [11]. MIMO channel configuration with 2 × 2 antenna pattern is shown in Fig. 1.

A technique is utilized for this type of antenna signal correlation in MIMO channel model for Rayleigh flat-fading channels. That can be mathematically described through the popular Kronecker channel model [13–14]. This model is defined as [15],

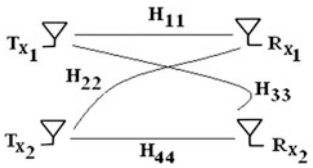


Fig. 1 MIMO channel configuration with different path gain H_{11}, H_{22}, H_{33} and H_{44} [12]

$$H_{Kron} = \frac{1}{\sqrt{tr[N_{H,R}]}} \left[\sqrt{(N_{H,R})}\right] S \left[\sqrt{(N_{H,T})}\right]^T \tag{3}$$

where, S is stochastic $n_r \times n_t$ matrix with independent identically distributed (i.i.d.) complex Gaussian zero mean unit variance elements. $N_{H,T} = E\{H^T H^*\}$ is correlation matrix at transmitting side, H^T is transpose matrix and H^* is conjugate matrix, $N_{H,R} = E\{HH^H\}$ is correlation matrix at receiving side. The significance of Kronecker model is involved with the transmitter, that does not affect the spatial characteristics of received signal at all. Now, another mathematical parameter of Kronecker product [16] of spatial correlation between receiver and transmitter is described as,

$$\begin{aligned} N_{MIMO} &= N_{H,R} \otimes N_{H,T} \left\{ \text{(for uplink)} \right. \\ &= N_{H,T} \otimes N_{H,R} \left\{ \text{(for downlink)} \right. \end{aligned} \tag{4}$$

where, $\otimes$ indicates Kronecker product.

Kronecker model is grown to be a popular model because of its simplicity in analysis. On the other hand, there is a big drawback in this model, because it forces both the link ends to be separable [17], irrespective of whether the channel sustains this condition or not. So, there is another model, Weichselberger is used to settle down the separability constraint of the Kronecker model and to permit for any arbitrary coupling between transmit and receive eigen base, i.e. to model the correlation characteristics at receiver and transmitter jointly.

Now, eigen value decomposition of receive and transmit correlation matrices can be introduced in the following way

$$\begin{aligned} N_{H,R} &= W_R \Lambda_R W_R^H \\ N_{H,T} &= W_T \Lambda_T W_T^H \end{aligned} \tag{5}$$

Weichselberger [18, Chap. 6.4.3] proposed

$$H_{weich} = W_R (\Omega^{\bullet}_{weich} \odot S) W_T^T \tag{6}$$

where, S is an i.i.d. complex Gaussian random fading matrix, $\Omega^{\bullet}_{weich}$ is defined as coupling matrix because its coefficients indicate how much power is coupled from

one transmitting eigenmode to another receiving eigenmode on average [18]. W_R and W_T are one-sided eigenbases and ⊙ defines Schur–Hadamard (element-wise) product.

In the previous two models MIMO channel model can be defined in the eigen space. Now another model virtual channel representation (VCR) is used to represent MIMO channel in beamspace. Particularly, the eigen-vectors are replaced by fixed and predefined steering vectors [19]. VCR can be mathematically defined as

$$H_{virtual} = G_R(\Omega^{\bullet}_{virtual} \odot S)G_T^T \tag{7}$$

where, orthonormal response and steering vectors represent columns of unitary response and steering matrix G_R and G_T. $\Omega^{\bullet}_{virtual}$ is defined as coupling matrix of VCR model.

4 Simulation and Experimental Results

This work is done on the basis of two different comparative reports. First comparison is done on the basis of capacity analysis of different array pattern of MIMO model and the next comparison is described on the basis of PDF of elements in SVD decomposition using same array patterns. In this paper, different set of parameters are used to execute the simulation. Simulation is done through MATLAB software and validated through Qualnet 6.1 simulation software [20]. Different antenna array patterns have been considered to perform the simulation. Set of numerical values of simulation parameters are shown in Table 1.

From Fig. 2 it is observed that higher capacity 21.5 bits/s/Hz is achieved at low level of correlation 4 × 4; whereas, lower capacity 7 bits/s/Hz is observed at high level of correlation 1 × 1.

Table 1 List of necessary simulation parameters

Parameters	Value
SNR variation	−10 to 20 dB
Different iterative values of simulation to observe the PDF of elements in SVD	1,000; 10,000 and 1,00,000
Types of MIMO antenna array pattern	1 × 1, 2 × 2, 2 × 3, 3 × 2, 3 × 4, 4 × 3, 4 × 4

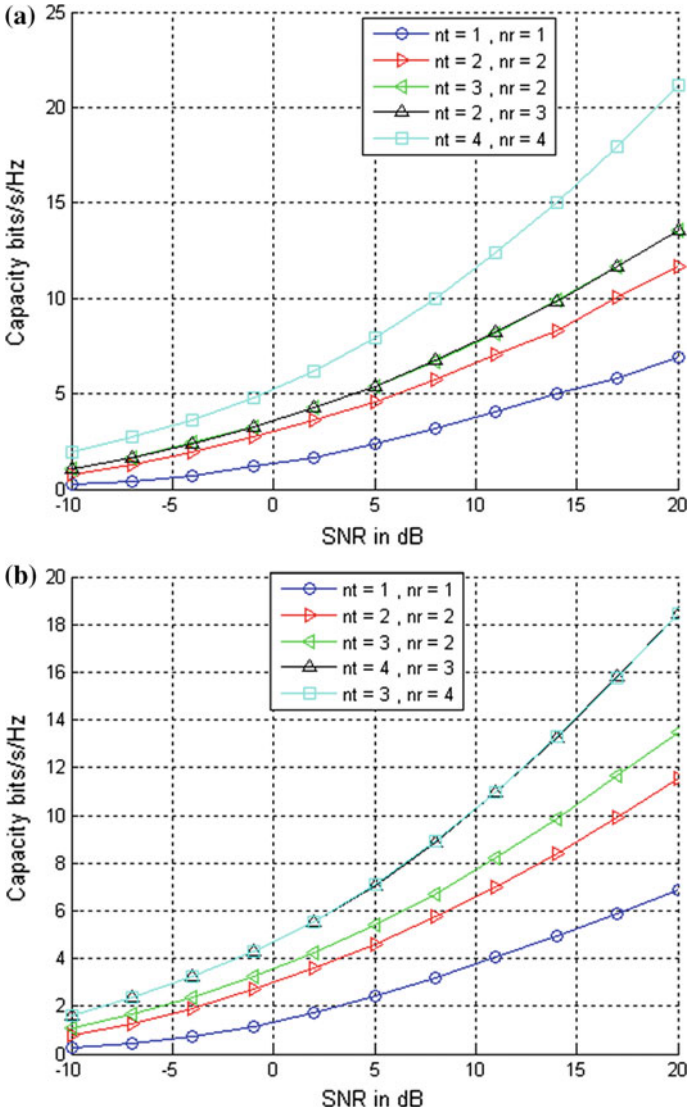


Fig. 2 **a**, **b** Capacity analysis for different array patterns at transmitter and receiver of MIMO

From the comparative analysis of Figs. 3, 4 and 5, it is shown that the random variation is lower at higher level of iterative values and PDF increases for lower level correlation of antenna array. Moreover, Lowest PDF is observed at higher level of correlation.

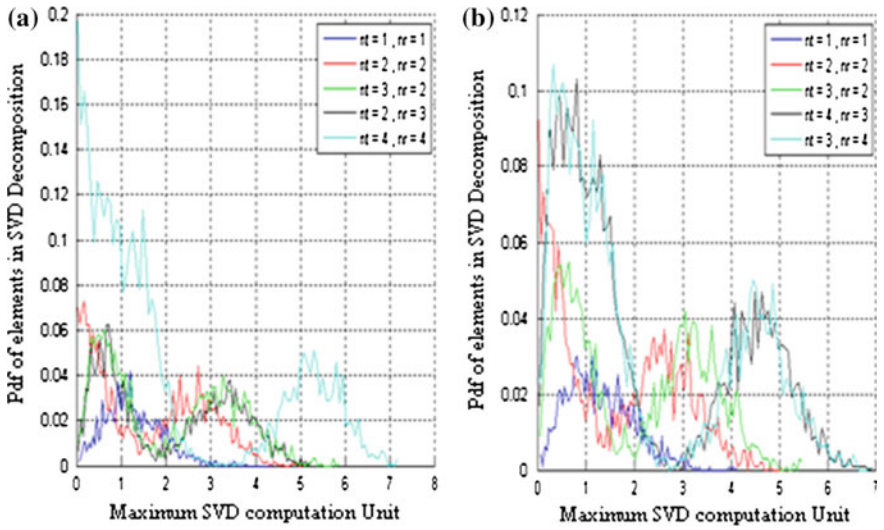


Fig. 3 **a**, **b** PDF of elements for 2 different array combinations in SVD decomposition for 1,000 iterative values of simulation

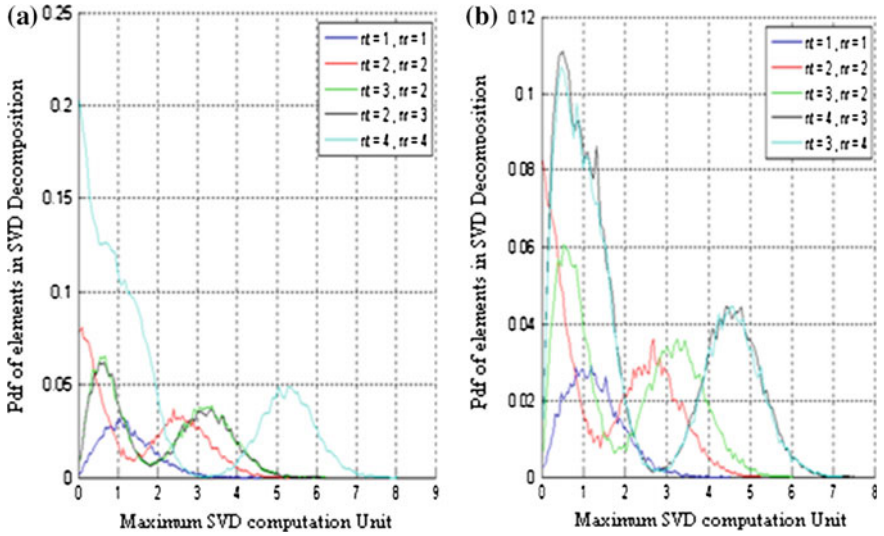


Fig. 4 **a**, **b** PDF of elements for 2 different array combinations in SVD decomposition for 10,000 iterative values of simulation

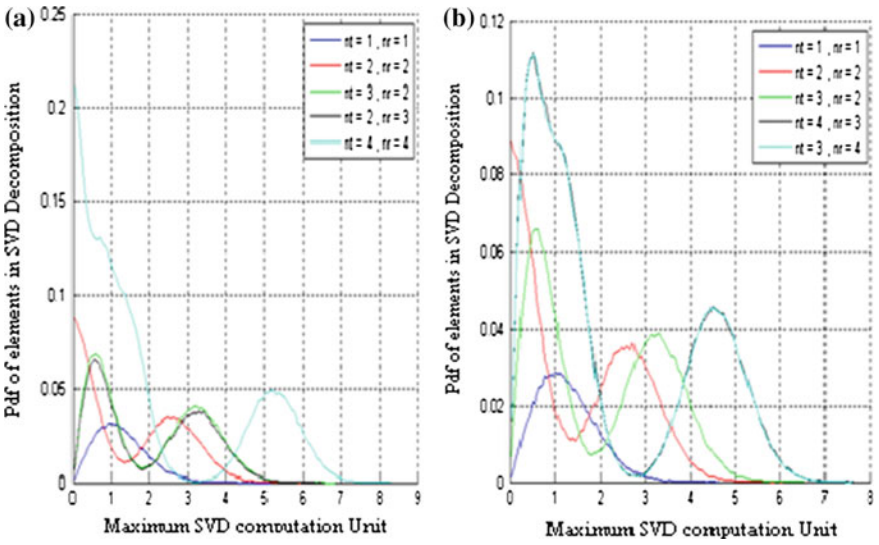


Fig. 5 **a**, **b** PDF of elements for 2 different array combinations in SVD decomposition for 1,00,000 iterative values of simulation

5 Conclusion

Main objective of this paper is to show a comparative evaluation of the distribution of different array patterns. Output of the system performance is produced for different iterative values to show the distribution of different array pattern in SVD decomposition. This work is verified through the compatible value of the parameters through Qualnet Software. Finally, it is concluded that better performance based on the higher capacity is possible to achieve through lower level of correlation.

References

1. Rusek, F., Persson, D., Lau, B.K., Larsson, E.G., Marzetta, T.L., Edfors, O., Tufvesson, F.: Scaling up MIMO: opportunities and challenges with very large arrays. IEEE Signal Process. Mag. **30**(1), 40–60 (2013)
2. Telatar, E.: Capacity of multi-antenna gaussian channels. Eur. Trans. Telecommun. **10**(6), 585–595 (1999)
3. Foschini, G.J., Gans, M.J.: On limits of wireless communications in a fading environment when using multiple antennas. Wireless Pers. Commun. **6**(3), 311–335 (1998)
4. Moustakas, A.L., Simon, S.H., Sengupta, A.M.: MIMO capacity through correlated channels in the presence of correlated interferers and noise: a (not so) large N analysis. IEEE Trans. Inf. Theory **49**(10), 2545–2561 (2003)

5. Jiang, Y., Li, J., Hager, W.W.: Uniform channel decomposition for MIMO communications. IEEE Trans. Signal Process. **53**(11), 4283–4294 (2005)
6. Alamouti, S.M.: A simple transmit diversity technique for wireless communications. IEEE J. Sel. Areas Commun. **16**(8), 1451–1458 (1998)
7. Tarokh, V., Seshadri, N., Calderbank, A.R.: Space-time codes for high data rate wireless communication: performance criterion and code construction. IEEE Trans. Inf. Theory **44**(2), 744–765 (1998)
8. Tarokh, V., Jafarkhani, H., Calderbank, A.R.: Space-time block codes from orthogonal designs. IEEE Trans. Inf. Theory **45**(5), 1456–1467 (1999)
9. Foschini, G.J., Chizhik, D., Gans, M.J., Papadias, C., Valenzuela, R.A.: Analysis and performance of some basic space-time architectures. IEEE J. Sel. Areas Commun. **21**(3), 303–320 (2003)
10. Golub, G.H., Van Loan, C.F.: Matrix Computations. Johns Hopkins UP (1996)
11. 3GPP Correlation properties of SCM: SCM-127, SCM Conference Call (2003)
12. Ghosh, S.: Performance evaluation with a comparative analysis of mimo channel on the basis of doppler shift and other probabilistic parameters in fading environment. IJMNCT (2014)
13. Shiu, D.S., Foschini, G.J., Gans, M.J., Kahn, J.M.: Fading correlation and its effect on the capacity of multielement antenna systems. IEEE Trans. Commun. **48**(3), 502–513 (2000)
14. Werner, K., Jansson, M., Stoica, P.: On estimation of covariance matrices with Kronecker product structure. IEEE Trans. Signal Process. **56**(2), 478–491 (2008)
15. Weichselberger, W., Herdin, M., Ozcelik, H., Bonek, E.: A stochastic MIMO channel model with joint correlation of both link ends. IEEE Trans. Wireless Commun. **5**(1), 90–100 (2006)
16. Kermoal, J.P., Schumacher, L., Pedersen, K.I., Mogensen, P.E., Frederiksen, F.: A stochastic MIMO radio channel model with experimental validation. IEEE J. Sel. Areas Commun. **20**(6), 1211–1226 (2002)
17. Özcelik, H., Herdin, M., Weichselberger, W., Wallace, J., Bonek, E.: Deficiencies of 'Kronecker' MIMO radio channel model (2003)
18. Weichselberger, W.: Spatial structure of multiple antenna radio channels. Institut für Nachrichtentechnik und Hochfrequenztechnik. (2003) Dec
19. Sayeed, A.M.: Deconstructing multiantenna fading channels. IEEE Trans. Signal Process. **50**(10), 2563–2579 (2002)
20. Qualnet Simulator: Scalable Network Technologies (2012). http://www.qualnet.com

Server Utilization-Based Smart Temperature Monitoring System for Cloud Data Center

Sudipta Sahana, Rajesh Bose and Debabrata Sarddar

Abstract The rise in demand for cloud computing services has thrown sharply into focus the subject of energy efficiency and cooling methods. The words "green" and "computing" can often translate into commercial and production successes, vendors and consumers alike are keen to optimize the services offered through cloud data centers as much as possible. While various existing methods help in bringing down rising temperatures of servers operating in cloud data center infrastructure, most authors would agree that pushing in cold air requires energy to be fed to cooling equipment and the associated infrastructure. Based upon existing research conducted, we approached the problem in a new light—concentrating on server utilization to regulate the temperature. We introduce Mean Utilization Factor concept that allows detecting and regulating the amount of cool air that is to be channeled in and around the servers within a cloud data center to bring down the operating temperature.

Keywords Temperature monitoring · Power consumption · Server utilization · Mean utilization factor · Cloud data center

S. Sahana (✉)
Department of Computer Science and Engineering, JIS College of Engineering, Kalyani, Nadia, West Bengal, India
e-mail: ss.jisce@gmail.com

R. Bose · D. Sarddar
Department of Computer Science and Engineering, University of Kalyani, Kalyani, Nadia, West Bengal, India
e-mail: bose.raj00028@gmail.com

D. Sarddar
e-mail: dsarddar@rediffmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_30

1 Introduction

The general consensus among experts is that maybe in the world, within the next century, production costs might outstrip the selling price of fossil fuels such as coal and oil; thus, stymieing efforts by major producing nations and corporations are required to ensure such supply that can be profitably mined or drilled, and supplied to consumers worldwide at sustainable rates. In a paper titled "Review of underground coal gasification technologies and carbon capture" the authors Stuart J Self et al. [1], discuss methods in coal production procedures to extend the available coal reserves to last several decades at least over and above the estimated 150 years at which no more coal can be extracted using the currently followed production practices. Although, nuclear power is a credible source, the high cost of maintenance, security and safeguarding of such power plants do not make it an entirely safe and bankable proposition in the eyes of some of the world's leading nations. In Japan, the catastrophe brought upon the Fukushima nuclear power plant by tidal waves generated by one of the largest tsunamis ever recorded in the history of mankind, closed down production for an indefinite period of time. Solar energy and wind farms have still some distance to go when compared to fossil fuels and nuclear energy which contribute to almost all of the world's energy needs. Over the years since it was realized that fossil fuels cannot last indefinitely, the focus has been to make energy efficient devices, machinery, consumer electronics, cars, etc.

While it has been found that a lion's share of the total energy used in a data center is used by servers, no less than 23% of energy is consumed by cooling infrastructure, as detailed in the paper "Cloud computing and its interest in saving energy: the use case of a private cloud" published by the authors Robert Basmadjian et al. [2] The authors present a proposed model in which a federation of several private and public clouds can be used to save around 20% of existing energy used. In presenting their work, they introduce energy optimization policies which are guided by models that predict power consumption. The underlying structure of the mechanism proposed to save energy comprises a three-step control loop. The steps involved are optimization, reconfiguration, and monitoring. The authors propose that the state of the data center be monitored continuously in order to achieve accuracy in resolving issues that involve energy consumption at various stages of the data center operation.

In a research article captioned "Cooperative Expendable Micro-Slice Servers (CEMS): Low Cost, Low Power Servers for Internet-Scale Services" authored by James Hamilton [3] of Amazon Web Services points out cooling as being the second most significant cost, only after server acquisition and before costs involved in distribution of power, and acquisition of power itself. While explaining the power used up in cooling the facility and distribution of power (difference between Total Facility Power and IT Equipment Power), the author familiarizes the reader with the Green Grid terms: Power Usage Effectiveness (PUE) and Data Center Infrastructure Effectiveness (DCIE). Here, the author has drawn reference to Green Grid. The Green Grid Power Efficiency Metrics: PUE & DCiE,

http://www.thegreengrid.org/gg_content/TGG_Data_Center_Power_Efficiency_Metrics_PUE_and_DCiE.pdf, 2007 [4].

We have organized the paper in several sections. We discuss the related research in the section titled "Related Work." In the subsequent section titled "Proposed Approach" we analyze our proposed method. The proposed algorithm and flowchart make up that follow next. In the section titled "Result Analysis," we discuss the results obtained through examination of our proposed model. We conclude our research and summarize our findings in the section captioned "Conclusion."

2 Related Work

In their work titled "Towards bandwidth guaranteed energy efficient data center networking" the authors Ting Wang et al. [5] explain their proposed solution for optimization of energy and making servers and allied infrastructure in a data center environment more energy-aware. In their work the authors describe wastage of energy as a result of not achieving peak capacity. Their work aims to address the situation by constructing a solution that can be described as energy proportional data center network without leaving any adverse effect on the throughput, nor undermining fault tolerance. Two schemes have been proposed for resource allocation, routing and scheduling of flow. Although the researchers point out the power distribution and cooling accounts for over 70% of overall power in comparison to a mere 15% consumed by the network infrastructure, the cooling factor achieve as a result of reduced power consumption by network is not explained. This, in our opinion, may be viewed as a shortcoming of this paper.

The authors, Xiaodong Wang et al. [6] in their work "Towards Optimal Sensor Placement for Hot Server Detection in Data Centers" introduce the problem of thermal monitoring and the absence of any efficient form of cooling operations in usual data center setups. Sans a systematic monitoring of temperature conditions with pinpoint accuracy as to the locations or spots within a data center, cooling systems often employ a holistic approach in lowering the temperature of an entire data center room without taking into account a strategy that necessarily looks into the micro aspect of rise in temperatures that correspond to a server running at near full utilization. This, therefore, leads to energy consumption levels that are unnecessary and could have been avoided. Sensor network technology has been already included in most data center thermal monitoring setups. Partly owing to its non-intrusive features, and more significantly owing to its robust nature to respond almost instantly to CPU or disk-related changes and activities. The authors point out that although such sensor network technology is an advantage to have in data center setups, most solutions tend to deploy sensors in a manner that seldom takes into account the temperature properties in a typical data center. The authors have worked around this problem of sensor placement by invoking a scheme that is based on Computational Fluid Dynamics (CFD). Using this CFD analysis in more than scenario involving situations where servers are overheating, the authors have

proposed applying of optimization techniques that would allow determination of placement of sensors in a manner such that detection of servers running hot can easily be detected. The authors have opined, based on empirical results, that their proposed model is effective and more accurate than general placement solutions which do not offer similar degree of probability in terms of detection of hot zones.

Computation Fluid Dynamics [CFD] has also been adopted by the authors Atsushi Itou et al., in their work titled "High performance parallel computing for Computational Fluid Dynamics (CFD)" [7] to solve temperature equations involving computation of the supersonic aerodynamic coefficient of a sphere representing a broken segment of a structure exposed to the effects of an explosion.

Using a robot to implement an autonomously conducted thermal and humidity map of a hitherto uncharted data center was mooted in the paper "A Robot as Mobile Sensor and Agent in Data Center Energy Management" by Hoi Chan et al. [8]. The authors propose deploying a robot as a form of mobile sensor which keeps uploading the data collected into IBM Tivoli Monitoring for Energy Management system for further processing and analysis. The authors state that although using a robot was not new, the proposed feature of a robot being aware of the thermal conditions in its surroundings, combined with automated diagnostic capabilities, is a novel concept. The authors propound using the robot as a mobile thermal sensor, in addition to deploying sensors at scattered locations. In their view, deployment such as this would allow data center administrators to quickly identify and isolate potential hot spots even before chance of an excessive overheat situation matures. However, it is not apparent how the authors propose to quickly detect workload changes in servers which bring about a change in the thermal characteristics. Further, chances are relatively higher in such a situation where static sensors are not placed close to certain locations where temperatures can rise to levels without immediate detection either by the robot or through any supporting mechanisms.

Computational Fluid Dynamics also form the basis of a thermal modeling tool, termed as "ThermoStat", by the authors Jeonghwan Choi et al. [9] of the paper titled "Modeling and Managing Thermal Profiles of Rack-mounted Servers with ThermoStat". The authors have presented the results of conducting several experiments with their thermal modeling tool to demonstrate how it can aid in designing solutions that can effectively manage thermal properties. The authors validate their findings with more than 30 temperature sensor measurements at assorted points in a 20 node server system that is mounted on a rack. Each of the servers that were mounted on the rack came equipped with dual Xeon processors, to effectively simulate production environment and obtain more accurate results through this tool. The findings of the authors in their work are most interesting; especially, in view of the fact, that their experiments conducted demonstrate that servers placed at the top and toward the rear in a rack tend to be hotter in a configuration where cold air is being blown in from the front and hot air extracted from the rear. Consequently, the authors correctly conclude placement of servers running on considerably higher loads to be placed near the bottom to take advantage of the cold air. Thus, a thermal profile with temperature-aware and load balancing management can be designed.

In their work "Resource Monitoring and Management with OVIS to Enable HPC in Cloud Computing Environments" the authors Jim Brandt et al. [10] discuss their proposal of using the OVIS project to monitor and manage resources for HPC applications in a cloud environment. The OVIS project is a system that is modular and has been designed specifically keeping in view of HPC metrics. What is significant about this tool for the purposes of our research is OVIS project's light-weight transport layer—The Lightweight Distributed Metric Service (LDMS)—that allows for collection and display of a variety of data, including those directly relating to temperature, fan speed, and CPU utilization. The former and the latter variables are of utmost importance in our own research and the method that we propose in this paper.

In identifying the need to optimize energy consumption, the authors, Gregory Katsaros et al. [11] in their research captioned "A service framework for energy-aware monitoring and VM management in Clouds" have presented a service framework that is able to deal with several important metrics all at once. Their proposed model lays the foundation for a virtual machine management system by monitoring energy consumption, calculating the energy efficiency, etc.. of a cloud infrastructure. The authors have been able to demonstrate positive results insofar as energy efficiency levels are concerned and have posted a 14% increase in infrastructural energy efficiency through the implementation of their derived virtual machine management strategy. As evident in this report compiled from the empirical results posted by the authors, energy-aware Infrastructure-as-a-Service platforms can be made more efficient. However, the authors do not highlight the factor of cooling which is a predominant issue in the present context where cloud computing has fast gained popularity.

Observing the need for load balancing to assure optimum use of available resources, the authors, Debabrata Sarddar et al., in their work titled "An Enhanced Cloud Network Load Balancing Approach Using Hierarchical Search Optimization Technique" [12] inferred that the absence of load balancing could cause a single node or more in a cloud infrastructure, to operate beyond levels that, in turn, lead to inefficiency and higher than usual rates of emission. The authors have, in this work, proposed using a model called Billboard Manager which computes certain set parameters to achieve load balancing. Although, the authors state that it is not a cost-effective approach, in our opinion this approach has its own merits which include optimization and distribution of processes among the nodes interconnected in a cloud setup.

While examining possible ways to reduce heat generation arising out of server overloads, the authors, Rajesh Bose et al., in their research report titled "An Energy Efficient Dynamic Schedule-based Server Load Balancing Approach for Cloud Data Center" [13] have introduced a method that requires following three different power consumption settings to correspond with a set utilization factor. The authors posit that through implementation of their proposed Data Center Manager (DCM) to trigger the appropriate power consumption setting, it would be possible for servers to have more power to operate instead of that power being diverted for additional cooling.

In their work "Smart temperature monitoring for data center energy efficiency", the authors, JunmeiQu et al. [14] discuss how to optimally place temperature sensors in data centers in a manner such that minutely controlling cool supply of air does not lead to setting off of temperature alarms simultaneously. The authors explain how they solve this problem through a two-step approach involving temperature snapshot that captures distribution of temperatures in a rack at a given point of time, and then extrapolating that information such that it is possible to compute sensor locations corresponding to the critical temperature zones. The authors affirm that by deploying sensors at these locations, cooling can be regulated efficiently, notwithstanding device workload fluctuations and changes, so that hot and cold spots can be obviated to a large extent.

3 Methodology

We approach the problem of increasing efficiency of a data center in a cloud infrastructure by designing a framework that would allow effective monitoring of the cooling system in place. As has already been understood from our literature survey, researchers have already uncovered evidence that efficiency of servers witness a steady decline with rising temperatures. Temperature rises may either be attributed to higher than standard processing requirements for extended periods of time, or owing to ambient conditions not directly attributable to server or device loads. Based on the studies conducted in the works [6, 9, 10], we introduce a unique framework that involves designing a control policy that regulates temperature by constantly monitoring sensor data. By switching off and turning back on power to the cooling system, we have tried to improve the efficiency and increase the life of the cooling equipment. Assisted by the research [10] and the features available, we propose to introduce a monitor that would record server utilization data in real-time, as well as regulate the cooling system. Further, our framework uses a dozen cooling units placed strategically within a data center based on studies conducted in the report [6]. However, at any given point, two of these cooling units will alternately switch on at intervals of 6 h even when all the servers remain idle. A dynamic schedule has been used that has been prepared based on data collected using procedures reported in the works [8–10]. We prepare a table that comprises of values computed as mean server utilization figures. The salient feature of this table is that we use previous studies conducted, to make a close approximation of the number of cooling units that are required to operate and maintain an optimum environment for any given range of mean server utilization values. The monitor is designed such that it always checks whether the mean server utilization factor exceeds the 95% mark. In case it detects possibility of this mark being exceeded as a result of accepting an incoming request, such request is kept in a buffer. The monitor checks the buffer for pending requests and mean server utilization factor for available servers at intervals of 5 min.

3.1 Proposed Algorithm

Our proposed algorithm broadly consists of five basic steps. These are as follows:

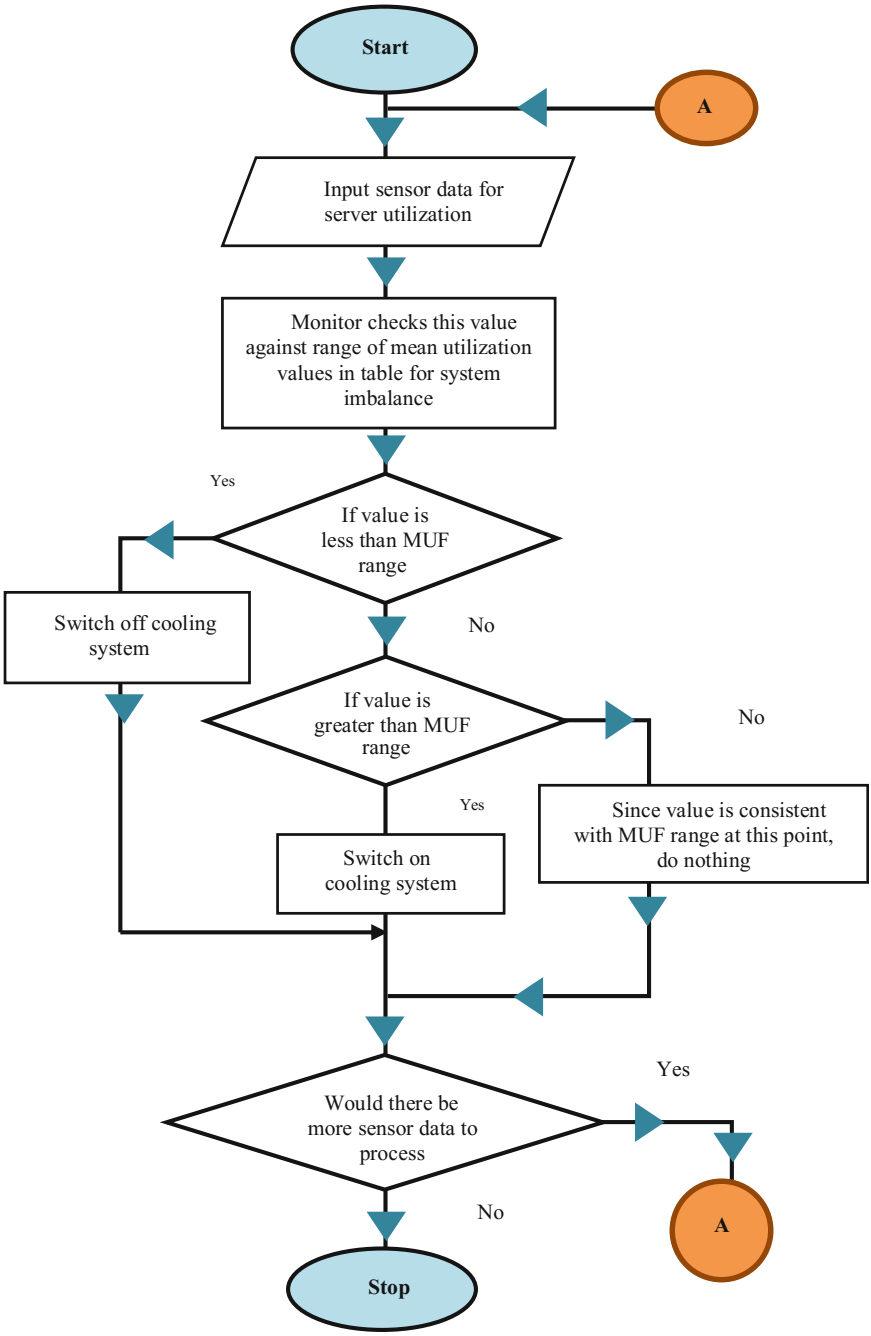
1. The Monitors gathers server data utilization values for a given number of servers for a set period. Usually this period can be a 24-h cycle during which these servers may be under observation for fluctuations in utilization subject to processing demands placed on it.
2. Based on the figures collected, a mean utilization factor (MUF) is calculated and a master table is prepared with corresponding values to indicate the number of cooling units that were required to keep the temperature at optimum levels at various MUF ranges for the given number of servers.
3. After the master table has been prepared, our proposed model kicks off in production mode. In this mode, the Monitor closely observes and keeps computing the real-time mean utilization factor of a server. If the number of running units corresponds to the observer MUF in accordance with the master table, the Monitor does nothing to alter the state of the cooling units.
4. In case the MUF detected is more, then the number of cooling units required to maintain an optimum state, according to the values in the master table, is turned on.
5. If the MUF computed is less, then the number of cooling units is turned off to reflect the number of cooling units required to be run continuously as per the master table.

The following is a master table which illustrates range of Mean Utilization Factors (MUFs) against the required number of cooling systems that are required to operate simultaneously. To obviate chances of a server freezing owing to excessive loads, the ceiling for utilization factor has been fixed at 95% (Table 1).

Table 1 Master table

Case number	Mean Utilization Factor (MUF)	Required cooling system
1	0 < X <= 10	2
2	10 < X <= 20	2
3	20 < X <= 30	3
4	30 < X <= 40	4
5	40 < X <= 50	5
6	50 < X <= 60	6
7	60 < X <= 70	7
8	70 < X <= 80	8
9	80 < X <= 90	9
10	90 < X <= 100	10

Flow Chart


Start
A
Input sensor data for server utilization
Monitor checks this value against range of mean utilization values in table for system imbalance
Yes
If value is less than MUF range
Switch off cooling system
No
If value is greater than MUF range
No
Yes
Switch on cooling system
Since value is consistent with MUF range at this point, do nothing
Would there be more sensor data to process
Yes
No
A
Stop

4 Result and Discussion

Application of our proposed method reduces unnecessary energy consumed by cooling systems. In actual practice, the trend is to have a fixed schedule for cooling systems to operate in a data center. In certain cases, there is no schedule at all. Consequently, there is no relationship with server utilization and the amount of cooling that is actually warranted to keep a thermal balance and, thereby, optimizing power consumption patterns. Indirectly, carbon emissions can also be controlled.

To illustrate the efficacy of our approach, a data center equipped with D number of servers each denoted by S; N number of cooling systems each denoted by C, consuming M watts per hour per C system is assumed.

In a general setup with no scheduling in place, the total energy consumed by all the cooling units in the data center is given as follows:

$$E_t = C_N \times M \times 24$$

where,

E_t is the total energy consumed, expressed in watts
C_N is the total number of cooling systems
M is the consumption rate of energy of one cooling unit in an hour
S_D is the total number of servers in states of either idle, running, or in sleep modes

In our proposed model, the total energy E_t consumed would not change if all S_D servers were to run at MUF of 95% or thereabouts. However, in the real world seldom do all servers in a data center operate near the peak MUF values all at once. Therefore, it can be said that where MUF values vary according to individual server load patterns, not all the cooling systems would be actively running throughout a given 24-h cycle. The total power E_p consumed, therefore, in such a scenario can be

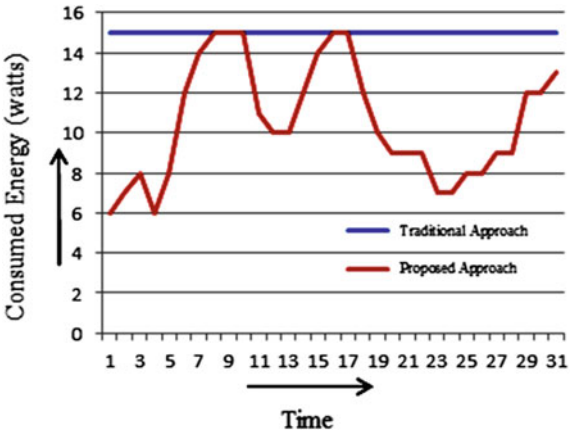


Fig. 1 Consumed energy versus time

said to be less than the total energy E_t consumed by all the cooling systems running actively throughout a 24-h cycle.

Thus, it is observed that since E_t is always greater than or equal to E_p, energy savings can be in terms of positive values (Fig. 1).

5 Conclusion

By conducting a relevant literature survey of existing research work, we have proposed through this paper, a smart temperature monitoring system which works off of server utilization factors. With the help of the research works conducted [6, 8–10], we are able to position our sensors in a manner such that blanketing of the entire data center with an unnecessary amount of sensors can be avoided. Thus, the cost of sensors is kept controlled within acceptable limits. The concept that we have proposed works on the principle that cool air would be kept flowing based on thermal patterns and heat signatures detected as a result of server workloads. In this manner, incidents of cold or hot spots can largely be minimized. As a welcome side-effect, reduction of power required for cooling implies reducing the carbon footprint of a data center to a significant extent. Further, life of the cooling equipment is extended as these no longer have to work actively non-stop and without the benefit of an active scheduling.

Future research work can improve upon our proposed approach using a MUF maximum threshold. Based upon the works [12, 13], load balancing and server-state switching can be implemented in conjunction with the report as presented in this paper.

References

1. Self, S.J., Reddy, B.V., Rosen, M.A.: Review of underground coal gasification technologies and carbon capture. Int. J. Energy Environ. Eng. **3**, 16 (2012). doi:10.1186/2251-6832-3-16
2. Basmadjian, R., De Meer, H., Lent, R., Giuliani, G.: Cloud computing and its interest in saving energy: the use case of a private cloud. J. Cloud Comput. Adv. Syst. Appl. **1**, 5 (2012). doi:10.1186/2192-113X-1-5
3. Hamilton, J.: Cooperative Expendable Micro-Slice Servers (CEMS): low cost, low power servers for internet-scale services. In: 4th biennial conference on innovative data systems research (CIDR), Asilomar, California, USA, 4–7 Jan 2009
4. The Green Grid Power Efficiency Metrics: PUE & DCiE (2007), http://www.thegreengrid.org/gg_content/TGG_Data_Center_Power_Efficiency_Metrics_PUE_and_DCiE.pdf
5. Wang, T., Qin, B., Su, Z., Xia, Y., Hamdi, M., Foufou, S., Hamila, R.: Towards bandwidth guaranteed energy efficient data center networking. J. Cloud Comput. **4**, 9 (2015). doi:10.1186/s13677-015-0035-7
6. Wang, X., Wang, X., Xing, G., Chen, J., Lin, C.-X., Chen, Y.: Towards optimal sensor placement for hot server detection in data centers. In: ICDCS '11 Proceedings of the 2011 31st International Conference on Distributed Computing Systems, pp. 899–908. ISBN: 978-0-7695-4364-2. doi:10.1109/ICDCS.2011.20

7. Itou, A., Nakanishi, T., Mizuguchi, T., Yoshida, M., Saburi, T.: High performance parallel computing for Computational Fluid Dynamics (CFD). In: Komatsu Technical Report, vol. 51 No. 156 (2005)
8. Chan, H., Connell, J., Isci, C., Kephart, J.O., Lenchner, J., Mansle, C., McIntosh, S.: A robot as mobile sensor and agent in data center energy management. In: Proceedings of the 8th International Conference on Autonomic Computing, ICAC 2011, Karlsruhe, Germany, 14–18 June 2011. doi:10.1145/1998582.1998610
9. Choi, J., Kim, Y., Sivasubramaniam, A., Srebric, J., Wang, Q., Lee, J.: Modeling and managing thermal profiles of rack-mounted servers with thermostat. IEEE Int. Symp. High-Perform. Comput. Archit. HPCA 205–215 (2007). doi:10.1109/HPCA.2007.346198
10. Brandt, J., Gentile, A., Mayo, J., Pebay, P., Roe, D., Thompson, D., Wong, M.: Resource monitoring and management with OVIS to enable HPC in cloud computing environments. In: IPDPS, 2009, International, Parallel and Distributed Processing Symposium, pp. 1–8 (2009). doi:10.1109/IPDPS.2009.5161234
11. Katsaros, G., Subirats, J., Oriol Fitó, J., Guitart, J., Gilet, P., Espling, D.: A service framework for energy-aware monitoring and VM management in clouds. Future Gener. Comput. Syst. **29**(8), 2077–2091 (2013)
12. Bose, R., Sahana, S., Sarddar, D.: An adaptive cloud service observation using billboard manager cloud monitoring tool. Int. J. Softw. Eng. Appl. **9**(7), 159–170 (2015). ISSN: 1738-9984. doi:10.14257/ijseia.2015.9.7.17
13. Bose, R., Sahana, S., Sarddar, D.: An energy efficient dynamic schedule based server load balancing approach for cloud data center. Int. J. Future Gener. Commun. Netw. **8**(3), 123–136 (2015). ISSN: 2233-7857. doi:10.14257/ijfgcn.2015.8.3.12
14. Chaudhry, M.T., Ling, T.C., Manzoor, A., Hussain, S.A., Kim, J.: Smart temperature monitoring for data center energy efficiency. ACM Comput. Surv. (CSUR) **47**(3), Article No. 39 (2015). doi:10.1145/2678278

Part VI
Civil Engineering and Structural Design

Assessment of Uncontrolled Intersections Through Calibration of VISSIM for Indian Traffic Conditions

Suprabeet Datta

Abstract The target of this study is to build up a VISSIM simulation model and align it to find out volume-to-capacity proportions of turning movements at uncontrolled intersections under Indian mixed traffic conditions. Driver and traffic behavior related parameters are adjusted after examination of the field information. The microscopic simulation outputs of the calibrated VISSIM model are contrasted with the field capacity values assessed from gap acceptance. Classified Movement volumes gathered at a four-legged uncontrolled intersection in the state of Maharashtra is utilized. Driver Behavior parameters (car-following, lane changing and lateral driving) are initially decided for homogeneous movement having one of the six classifications of vehicles considered in the stream and after that the outcomes are accumulated to get the estimations of these parameters for a mixed stream. Every single other factors and movement activities on every methodology are kept as steady. The calibrated VISSIM model is then used to decide capacities (volumes) by flooding a section at once for non-priority movements. Simulated volumes were obtained for each movements and turns after the 7th run. Gap acceptance capacity is calculated using HCM 2010 formula for the intersection. Volume-to-capacity ratio is used as the critical measure for assessing traffic flow operations within the intersection. After assessment through calibration it was found that the uncontrolled intersection is operating under low to moderate congestion (volume-to-capacity ratio less than 0.85) thus experiencing lesser service delays.

Keywords Capacity · Calibration · Intersection · Simulation · Uncontrolled · Volume-to-capacity ratio

S. Datta (✉)
Department of Civil Engineering, JIS College of Engineering, Kalyani, Nadia 743145, West Bengal, India
e-mail: sdce90@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_31

1 Introduction

The convenient and mounting populace of creating countries like India impacts the travel example of a particular group from their beginning to any goal and along these lines upsurges the quantity of vehicles in the city upsetting the general transportation structure. This causes congestion in the midst of pinnacle hours. The movement and transportation facility in creating countries like India, China, Indonesia, etc., are not exactly the same as that in created countries like United States, United Kingdom, etc., as movement piece and level of road side exercises vary in examination. Vehicle modes in developing countries exhibit a limitless variability which makes the movement stream rather heterogeneous. This movement stream involves transport methods of contrasting element traits having the same road space. In the road system, intersections are the noteworthy foundations for bottlenecks causing congestion. An intersection is a traffic facility where two or more roads either meet or cross at appraisal. Unsignalized intersections involve noteworthy section of the street systems in India. On the other hand, traffic engineers and researchers are significantly more propelled in roundabouts, grade-separated, signalized intersections, and interchanges as a delayed consequence of which lesser thought is being given toward unsignalized intersections. Unsignalized intersections can be assembled as right side/left side need controlled, stop-controlled, yield-controlled, incompletely controlled, and uncontrolled sorts. Capacity and level of service assume critical gaps in assessment of an unsignalized intersection. "*Capacity represents the maximum sustainable hourly flow rate at which persons or vehicles reasonably can be expected to traverse a point or a uniform section of a lane or roadway during a given time period under prevailing roadway, environmental, traffic, and control conditions*." [1] Flow rate as per U.S. Highway Capacity Manual is defined as "*the equivalent hourly rate at which vehicles pass over a given point or section of a lane or roadway during a given time interval of less than 1 h, usually 15 min*" [2]. VISSIM is a microscopic, time step, and conduct based reproduction stage created to demonstrate urban movement and open travel operations. The system can examine activity and travel operations under imperatives, for example, path design, movement organization, movement signals, travel stops, and so forth, therefore making it a helpful device for the assessment of different choices taking into account transportation building and arranging measures of adequacy. In addition, perceptible factual examination of activity information, no doubt, will not be able to totally get the heterogeneity developing out of the blended movement conditions and way indiscipline existing on Indian roads and in this manner, microscopic simulation using VISSIM 7.0. The essential goal of this study is to simulate and validate an uncontrolled intersection in VISSIM to get critical movement rates (capacities) by flooding a section at once for non-need developments. The proposed research incorporates the means of (a) Selection of reasonable convergences; (b) Data Collection; (c) Calibration Parameter Identification; (d) Experimental Design of the Network in VISSIM; (e) Running the Simulation; (f) Finding out the most extreme streams (limits) for every development from the

reenactment model created; (g) Validate the reproduction comes about with the field extricated estimations of stream rates for transferability.

2 Review of Previous Literatures

Capacity for unsignalized crossing points is evaluated by different methodologies which can be described as deterministic and probabilistic. The principal strategy is the gap acceptance technique (GAP), which was created in Germany and is also used as a piece of convergence limit estimation at United States and in a couple of European Countries. The key tenet of this strategy is to appraise the limit at unsignalized intersections in light of so–called critical gaps and follow–up times for the vehicles from the minor road. The second strategy is the empirical regression technique, which is mainly engaged around, scrutinizes ascending from the United Kingdom [3]. This procedure is engaged around a broad number of field data in current British roads by the use of regression functions. This philosophy of capacity estimation also considers elements like street geometric diagram, visibility distances, demand flows, turning proportions, and vehicle classes. The third methodology in ascertaining limit of unsignalized intersections is the "Conflict Technique." This new approach is engaged around the idea of "Addition of critical movement flows" [4]. Various studies have been done to adjust and accept VISSIM model for freeways, arterial streets under Indian mixed traffic conditions, few of them are inspected here. Among these studies on calibration of VISSIM, in particularly, Fellendorf and Vortisch [5] provided a discussion of the car following and driver behavior logic that is incorporated in VISSIM. The authors discussed detailed analysis of the Wiedemann car following model implemented in the VISSIM. Park and Schneeberger [6] used Latin Hypercube sampling along with a linear regression model to generate scenarios related to traffic situations and solved to match the travel times in field and simulation. Studies identified with calibration of VISSIM for expressways includes [4] and [7]. Studies related to calibration of VISSIM for freeways include [8] and [9]. Mathew and Radhakrishnan [10] calibrated non-lane-based heterogeneous traffic flow at signalized intersections. The authors identified the calibration parameters by a sensitivity analysis and optimized parameter sets by minimizing the error between the field and simulated delay through a Genetic algorithm. In some similar studies like in [11] and [9] calibrated VISSIM for signalized intersections using multi parameter sensitivity analysis. The authors determined the optimum values for these parameters by minimizing the error between the simulated and field delay using a genetic algorithm. This time, multiple criteria were included in the optimization formulation through insertion of constraints. Viti et al. [8] proposed a method to calibrate the critical gap parameter in microscopic simulations focused around individual travel time perceptions, and a basic formulation for the count of critical gap parameters that considers the expanding danger taken by drivers on the minor road with expanding volumes on the significant street. All above studies said are connected either expressways or

signalized intersections or roundabouts. Satisfactory work has been done on estimation of capacity through calibration of microscopic simulated parameters at unsignalized intersections under mixed traffic conditions and hence this study aims to calibrate driver behavior/car-following and lane-changing parameters for the designed uncontrolled intersections in VISSIM.

3 Methodology Adopted for the Study

The method adopted in this study is data extensive. Detailed data regarding traffic parameters such as vehicle composition, speed, acceleration, deceleration, trajectories, lane changing, has been extracted. The traffic speed data extracted for the intersection is attached in appendix. The methodology to be taken care with respect to broader aspect of this study is sequenced through the flow-chart mechanism represented by Fig. 1.

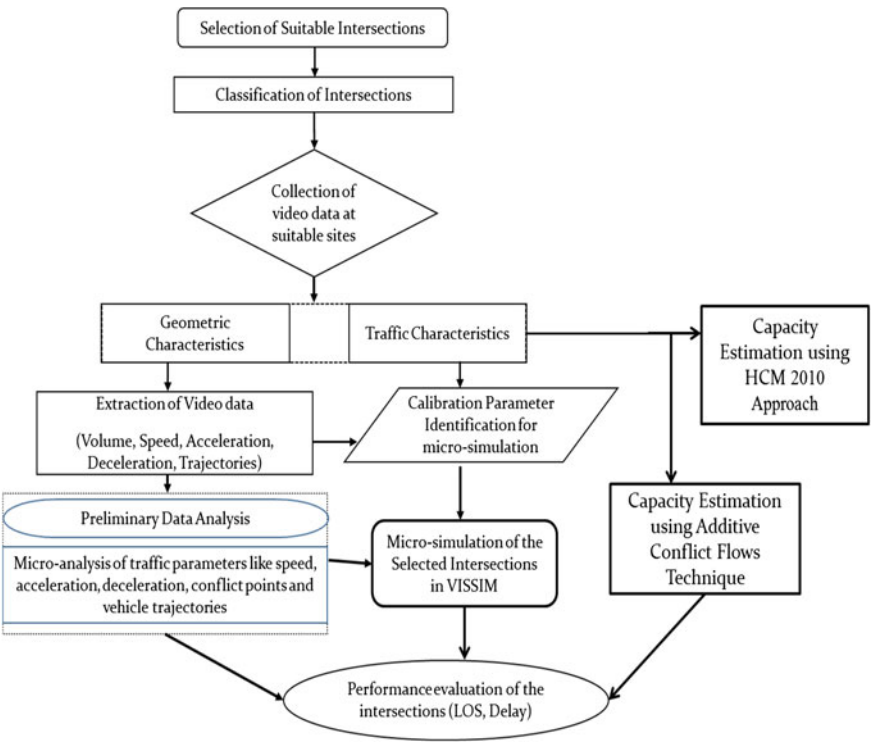


Fig. 1 Methodology adopted for broader aspect of the study

4 Data Collection and Extraction

Data collection and extraction is a vital undertaking before building up any model. The data extraction process was extremely exhaustive. Information for this study contains one four-legged uncontrolled intersection from the city of Kolhapur situated in the north-eastern province of Maharashtra, India. The intersection was used for gap acceptance studies in the past and thus was found to be suitable for capacity analysis. Figure 2 shows the detailed geometry of the intersection utilized for simulation as a part of this study. The major road is four-lane divided while the minor road is two-lane undivided. Although stop signs are present on the minor road but drivers stop only when major stream gaps are extremely less to accept. Drivers cross the intersection in two stages. Firstly, they cross the adjacent approach/lane and if essential wait on the median refuge area to cross the second conflicting movements prevalent on the second approach/lane.

The data extracted is utilized to discover different traffic characteristics at the intersection. From the visual observations, approximate influence area for an intersection is decided. Intersection influence area consists of major approach width, minor approach width, and the conflict area shared by both the minor and major approaches. Classified traffic volume data was obtained manually after playing the videos for one and a half hour in the laboratory during which all right-turning, left-turning and through traffic vehicles from each approach were considered. Since vehicle trajectories and speeds at different locations are needed, the pavement surface was marked with paint. The markings were divided at 5 m interims on all approaches starting from the centre of intersection. The markings encouraged the making of grids on the recorded video in the Traffic Data Extractor

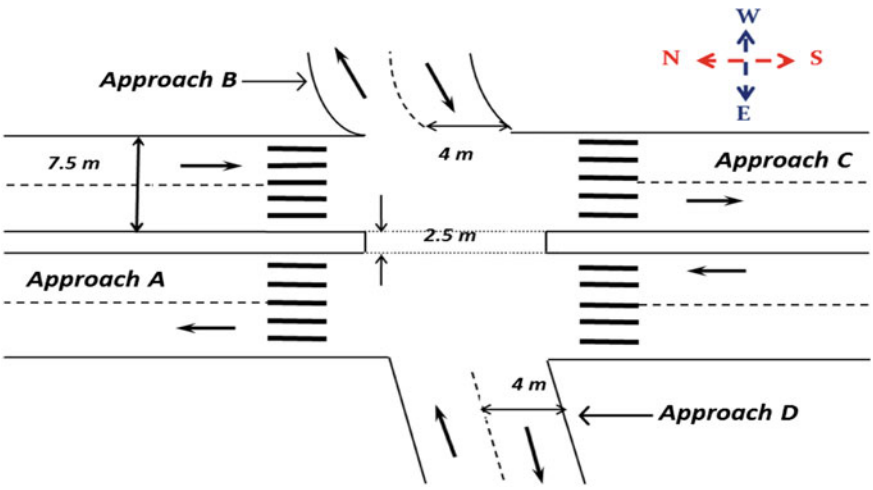


Fig. 2 Geometric details of the Kolhapur City four-legged uncontrolled intersection on national highway 204

Fig. 3 Vehicle trajectory extraction in Traffic Data Extractor (TDE)

(TDE). Traffic Data Extractor (TDE) is a system intended with the end goal of data extraction for the development of Indian Highway Capacity Manual (Indo-HCM) Project. This program contains tools for vehicle trajectory extraction, speed extraction and vehicle volume extraction from a video-based study. Speed information was acquired subsequent to playing the recordings on the screen and drawing fictitious lines at 20 m separation on the traffic data extractor interface. Vehicular trajectory information was extracted utilizing the same extractor as appeared in Fig. 3 after calibrating the recordings.

5 Preliminary Analysis of Video-Graphic Data

5.1 Traffic Composition and Speed Data Analysis

The classified traffic volume for different directional movements at the Kolhapur city intersection is shown in Table 1. There were a total of 14 directional movements (6 major and 8 minor) for the intersection. Out of these fourteen, classified traffic volumes of ten movements are shown in Table 1. The hourly traffic volume for Kolhapur city intersection was found to be varying from 125 vehicles per hour per lane during morning peak time to 435 vehicles per hour per lane at afternoon peak time. The total amount of right-turning vehicles during morning peak (9:07 to 10:51 AM) range from 588 to 1359 vehicles. Thus the hourly flow rate for right turns varied from 196 vehicles per hour to 453 vehicles per hour. The hourly flow rate for left turns ranged from 359 to 169 vehicles per hour which is lower than the

Table 1 Classified traffic volume at Kolhapur City four-legged intersection (for one and half hour data each for morning and afternoon)

Duration	Vehicular class	RT from approach C to D	RT from approach D to C	RT from approach B to A	RT from approach A to B	LT from approach D to A	LT from A to D	LT from approach B to C	LT from approach C to B	Approach A	Approach C
Morning peak (09:07–10:51)	Cars	346	68	94	110	91	80	83	103	273	284
	Two-wheelers	117	42	76	92	88	83	56	101	257	249
	Three-wheelers	54	37	35	31	44	29	38	26	33	40
	Sport utility vehicles	41	49	49	43	49	54	43	60	75	81
	Heavy vehicles including busses	69	31	23	31	63	27	54	42	32	71
Afternoon peak (14:03–15:35)	Cars	104	71	93	113	93	114	66	131	298	279
	Two-wheelers	87	53	27	68	47	63	44	58	88	83
	Three-wheelers	43	37	30	14	27	21	33	16	41	37
	Sport utility vehicles	78	30	32	48	58	52	58	36	60	73
	Heavy vehicles	56	25	17	42	29	33	39	12	56	48
Total		975	443	467	592	530	587	517	579	1213	1245

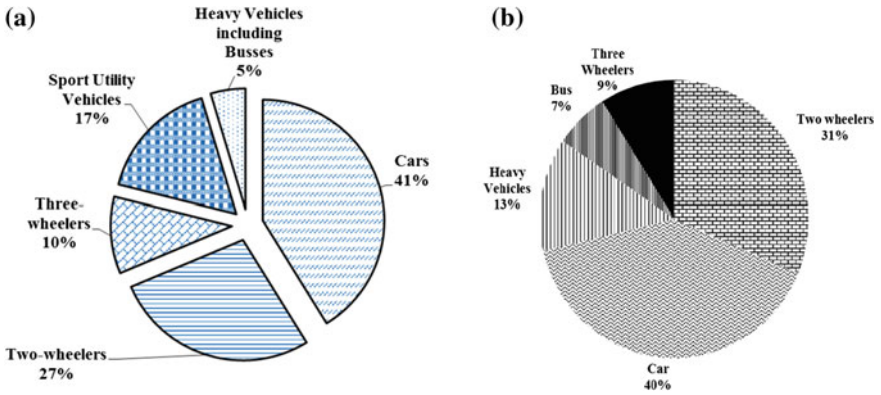


Fig. 4 Traffic composition for **a** Major approaches A & C; **b** Minor approaches B & D for the Kolhapur uncontrolled four-legged intersection

right- turning flow rate. The proportion of two-wheelers is highest in this intersection and is mainly because inhabitants living near to this area prefer two-wheelers for shorter trips and in areas of high congestion.

The values in Table 1 represents number of vehicles taking either right turns (RT) or left turns (LT) from major and minor approaches of the four-legged intersection. The traffic was sub-divided into five vehicular categories namely standard cars, two-wheelers, three-wheelers including auto-rickshaws, sport utility vehicles and heavy vehicles consisting of trucks and busses. Non-motorized transits such as bicycles and pedestrian compositions were found to be very less at the intersection and thus were neglected for simulation. Traffic composition of the major road approaches (A and C) for the Kolhapur city intersection is shown in Fig. 4. The share of two-wheelers is quite high (27%) with respect to three-wheelers (10%) as because the intersection is on a national highway (NH-204) and far away from conventional residential and commercial regions located inside the main city area. At national highways, cars and sport utility vehicles are mostly predominant and thus have greater composition (41% and 17%, respectively). Vehicle Speed data was also extracted after playing the videos in the laboratory for a period of 3 h. Figure 5a, b shows the frequency distribution plot for the minor road traffic (both approaches B and D) and major road traffic (both approaches A and C) for Kolhapur city four-legged intersection.

6 Microscopic Simulation

6.1 Introduction

Traffic simulation models have been broadly used to study traffic operations and systems impacts because simulation is more secure, less excessive, and faster than

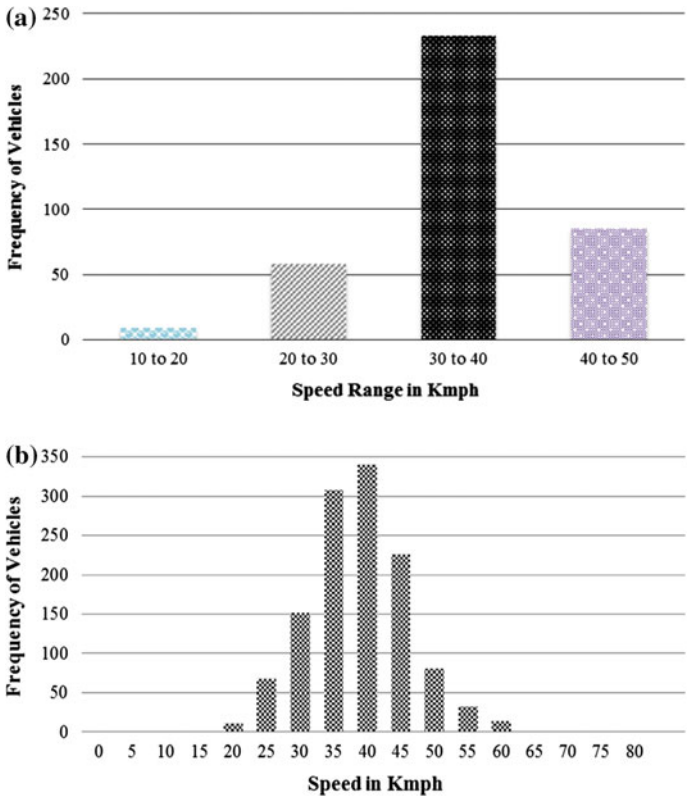


Fig. 5 Vehicle speed frequency distribution for **a** Minor approaches B & D; **b** Major approaches A & C for the Kolhapur uncontrolled four-legged intersection

field use and testing. "Microscopic simulation," sometimes called micro simulation, implies that every entity (auto, train, and individual) of reality that is to be simulated is simulated individually, in this way considering all pertinent properties. The same holds for the connections between the elements. The inverse would be a "macroscopic simulation," in which the description of reality is shifted from individuals to "averaged" variables like flow and density.

So as to utilize VISSIM reproduction model, first it needs to adjusted and accepted for existing traffic condition. Be that as it may, modifications to the default behavioral parameters are essential to effectively simulate Indian heterogeneous traffic conditions. Before calibrating gap acceptance parameters at uncontrolled intersection, we need to calibrate Indian heterogeneous traffic situation on road links. The traffic simulator in VISSIM is a microscopic traffic flow simulation model including the car following and lane change logic. VISSIM uses the psycho-physical driver behavior model developed by Wiedemann in 1974 and 1999.

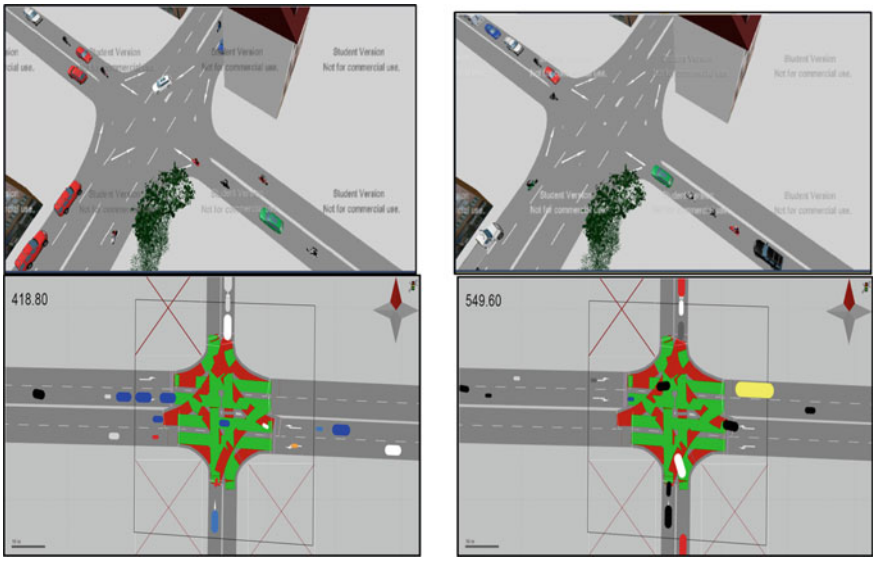


Fig. 6 Representation of the Kolhapur uncontrolled four-legged intersection in VISSIM 8.0 showing minor and major approach queuing before gap acceptance during simulation runs

The reason for visualization in a micro-simulation model is to approximately and somewhat accurately represent the field conditions. A model can't be esteemed calibrated if the animations are not practical. Animations from seven to eight runs were seen to recognize at which percentile they were not worthy. Figure 6 shows some praiseworthy simulation screenshots showing major and minor approach queuing before gap acceptance which was later validated from the video-graphic data for the Kolhapur intersection.

6.2 Identification of Calibration Parameters

Model calibration is the process of modifying and determining a set of model parameters, based on modeling judgment and data collected, that accurately represents the prevailing field conditions of a given study area. Both system and operational calibration are performed in VISSIM 8.0 which includes:

(1) System level calibration includes assumptions on vehicle route choice, traffic demand inputs, traffic compositions, study area boundaries, seeding period, and temporal distribution of demand and routing.
(2) Examples of operational calibration parameters include car following characteristics (headway, standstill distance, safety distance), lane changing accepted

deceleration rates, routing lane change distance, and lane selection. The operational calibration requires the most time and resources to complete.

Table 2 represents the all the vehicular dimensions, acceleration and deceleration values for all the six different types of modes fed in the simulation framework. Tables 3 and 4 lists down the Weidmann 99, Weidmann 74 (Driver Behaviour) and lateral driving and lane change parameters incorporated in the simulation model.

The lateral distance between two-wheelers, auto-rickshaw, car and heavy vehicles including bus were investigated from the video-graphic data and using the traffic data extractor software (TDE). Average standstill distance before gap acceptance, look-ahead distance, look-back distance, major stream headways, oscillation acceleration values were obtained from trajectory data.

Table 2 Vehicle dimensions, acceleration and deceleration characteristics feeded into the VISSIM simulation model for calibration

Sl. No.	Vehicle type	Actual size (m)		Acceleration (m/s^2)		Deceleration (m/s^2)		Des. speed (kmph)
		Width	*Length*	*Max.*	*Des.*	*Max.*	*Des.*	
1	Two-wheelers	0.6	1.8	2.5	1.7	1.7	0.2	48
2	Auto-rickshaw	1.4	2.6	0.9	1.1	1.1	0.8	36
3	Cars	1.7	4.7	1.5	1.2	1.2	1.0	35
4	SUVs	2.137	4.90	1.2	1.1	0.8	0.9	32
5	LCVs	2.326	5.411	0.8	0.8	1.1	0.8	27
6	Heavy vehicles	2.5	10.3	1.3	0.8	1.4	0.6	20

Table 3 Driver behavior related parameters and Weidmann 74 model

Parameters	Default values	Values incorporated
Driver behavior parameters		
Look-ahead distance (m)		
Minimum	0	10
Maximum	250	30
No. of observed vehicles	4	4
Look-back distance (m)		
Minimum	0	10
Maximum	150	50
Weidmann 74 car following characteristics		
Average standstill distance (m)	2	1–2
Additive part of safety distance	3	0.1–2
Multiplicative part of safety distance	3	0–3

Table 4 Lane changing and lateral driving parameters to be considered

Parameter	Default value	Values incorporated
Lane change		
Waiting time before diffusion (s)	60	120
Minimum headway (m)	0.5	0.11
Lateral behavior		
Desired position at free flow	Middle	Any
Diamond shaped queuing	No	Yes
Consider next turning direction	No	Yes
Overtake on same lane	No	On left and right
Min. lateral distance at 0 and 50 (m)		
Two-wheeler	1, 1	0.1, 0.4
Auto-rickshaw	1, 1	0.3, 0.5
Car	1, 1	0.5, 0.7
Heavy vehicles like bus and trucks	1, 1	0.7, 1.0

Table 5 Desired speed and acceleration values obtained from the simulation

Vehicle category	Desired speed (Kmph)	Desired accelerations (m/s^2)
Auto-rickshaws	29.75	1.15
Two-wheelers	40.87	2.31
Cars	32.89	1.35
LCV	25.53	0.677
SUV	29.24	0.697
Trucks	19.53	0.955
Bus	20.17	0.457

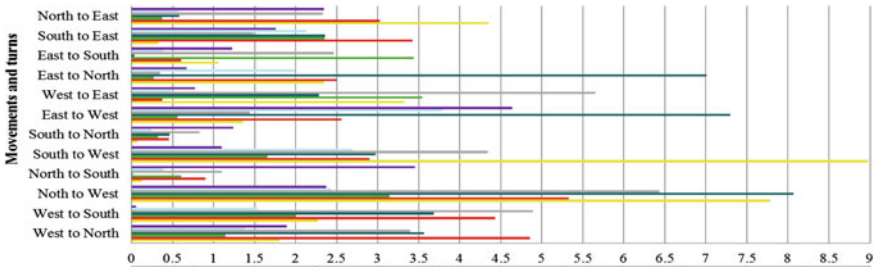
7 Results and Analysis

7.1 Data Collection and Node Evaluation Results from VISSIM

Table 5 lists down the desired speed and accelerations for seven different modes as obtained after 5th simulation run in VISSIM after calibration. The values were found to be close to that obtained after extraction of video data in TDE.

From Fig. 7 it is clear that stopped delay before gap acceptance is maximum (i.e., 8.98 s) for Light commercial vehicles (LCV) while taking major approach right turns (South-West).

Stopped delay for minor approach through movements (i.e., west-east and vice versa) are found to be maximum for auto-rickshaws (Auto) and busses whereas stopped delay for major approach left-turning movements (i.e., north-west and south-east) were found to minimum for auto-rickshaws (Auto) and maximum for heavy goods vehicles (HGV). The reason is the forced gap acceptance of



	West to North	West to South	Noth to West	North to South	South to West	South to North	East to West	West to East	East to North	East to South	South to East	North to East
Auto	1.89	0.05	2.37	3.45	1.1	1.24	4.64	0.77	0.67	1.22	1.76	2.34
Car	1.37	1.53	2.42	0.38	2.68	0.23	3.77	0.47	1.99	0.38	2.13	0.57
Bus	3.39	4.89	6.43	1.1	4.34	0.82	1.44	5.65	0.34	2.46	1.48	2.33
HGV	3.56	3.68	8.07	0.01	2.97	0.46	7.3	2.28	7.01	0.03	2.36	0.58
Bike	1.14	1.99	3.14	0.6	1.65	0.32	0.56	3.54	0.27	3.44	2.35	0.37
SUV	4.85	4.43	5.33	0.9	2.9	0.45	2.56	0.37	2.5	0.6	3.42	3.02
LCV	1.8	2.27	7.78	0.12	8.98	0.06	1.36	3.32	2.34	1.05	0.33	4.35

Stopped Delay in seconds for 3600 simulation seconds run

Fig. 7 Stopped delay comparison for different vehicle types obtained from the simulation runs (after 7 runs)

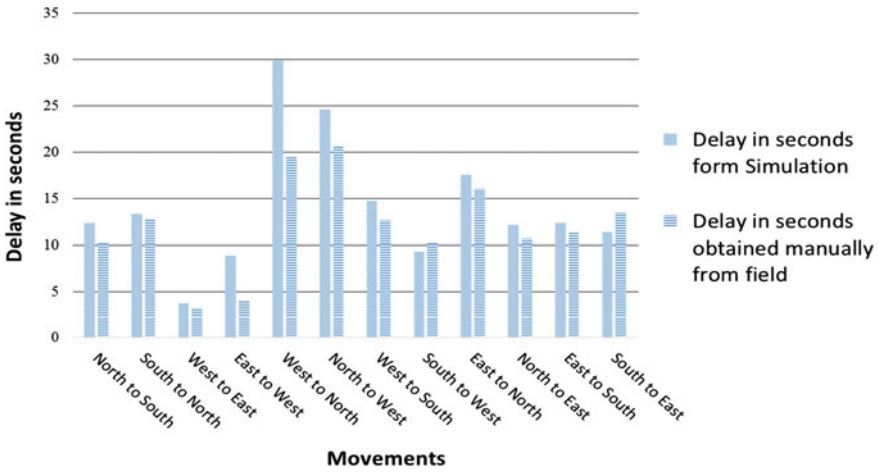


Fig. 8 Comparison of service delay from simulation and video data for different movements (after 7 simulation runs)

auto-rickshaw drivers during left turns resulting in increase of waiting time for heavy goods vehicles (Trucks) at the intersection which was validated from the video-graphic data.

After comparison of service delay manually obtained from the traffic data extractor (TDE) and that obtained after the 7th simulation run it can be inferred that the simulation model is an almost correct representation of the field traffic flow conditions as validated from Fig. 8.

Queue length for major approach (Approach A), i.e., traffic emerging from south-wards is maximum (i.e., about 40.45 m up to back of queue) which results in maximum stopped delay for south-west turning movement. Figure 9 shows exponential increase in minor approach (both B and D) service delay with increase in minor approach queue length which is common phenomenon observed during simulation.

Table 6 enumerates the values of density (in veh/km), speed (km/hr), volume (vehicles per hour) obtained after the 7th simulation run in VISSIM. The gap

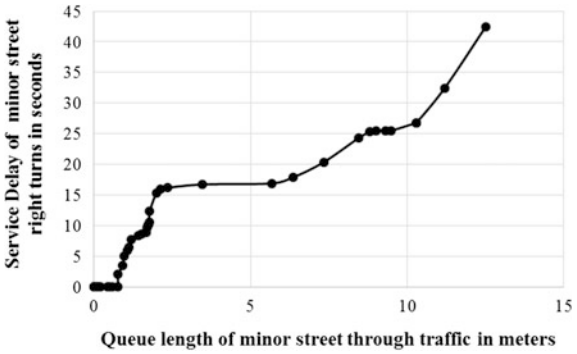


Fig. 9 Service delay of major street right turns versus queue length plot obtained from simulation results

Table 6 Density, speed and volume of the links obtained from the simulation runs (7nos.) for morning peak hour (9:07–10:07 AM)

Links	Density in veh/km from simulation	Speed in Kmph from simulation	Volume in veh/hr from simulation	Field capacity in veh/hr as per HCM 2010 formula	Volume-to-capacity ratio
West-App C	41.5	21.8	771.7	3400.5	0.227
West-App A	15.7	32.9	515.4	2607.2	0.198
Toward North-App B	10.4	32.9	342.6	988.1	0.347
Toward South-App B	41.2	10.5	398.4	1100.8	0.362
East-App C	24.5	32.1	787.0	2508.8	0.314
East-App A	33.4	23.0	756.0	1292.6	0.585
Toward North-App D	16.1	19.7	297.8	1045.0	0.285
Toward South-App D	17.8	31.7	566.2	1281.7	0.442

acceptance capacity values, as per the Highway Capacity Manual (HCM) [1], are also found out and the volume-to-capacity ratios were computed as shown above. After assessment of the volume-to-capacity ratio values (<0.85) with reference to HCM [1] it can be concluded that the intersection is operating under capacity and thus excessive delays were not experienced and the intersection operated under moderate to low congestion corresponding every movements and turns as validated from Table 6.

8 Summary and Conclusions

This study deals with the selection of intersections and the data extraction procedure involved. For this purpose an uncontrolled intersection from the city of Kolhapur on the National Highway 204 is selected. It also emphases on the capacity analysis through micro simulation of Kolhapur city four-legged uncontrolled intersection after calibrating the model with driver behavior (Weidmann 74 Car following), lane changing, and vehicular characteristic parameters. All these parameters were obtained from the video data extracted in the laboratory. Calibration includes modification of VISSIM Weidmann 74 car following behavior and lane-changing parameters. Flows and turning movements were also used as traffic input parameters for the experimental design of the VISSIM simulation model. Other parameters like accelerations, deccelerations, maximum speeds, minimum speeds, desired speeds, and vehicular dimensions for different vehicle types were calibrated based on the field data.

After simulating the intersection data, the minor road through traffic volume for both the simulated and field extracted data (actual field minor road through traffic flows) were compared. Variation of service or vehicle delay for minor road right-turning vehicles with respect to major road through traffic volume are shown for the simulation outputs. The study also shows the variations of minor road queue length with gradual increase in major road through flow as obtained from simulation outputs. After assessment of the volume-to-capacity ratio values (with reference to [1] it can be concluded that the intersection is operating under capacity (i.e., v/c ratio less than 0.85) and thus excessive delays were not experienced and the intersection was thus operated under moderate to low congestion.

Acknowledgements This study has been conducted as a part of research works for academic benefits and the author would like to acknowledge and thank Prof. Gopal R. Patil (Associate Professor, IIT Bombay) for providing the video-graphic data required to complete this study.

References

1. Highway Capacity Manual: Transportation Research Board of the National Academics, Washington, D.C (2010)
2. Highway Capacity Manual: TRB. National Research Council, SR-204, Washington D.C. (1994)
3. Kimber, R., Coombe, R.D.: The Traffic Capacity of Major/minor Priority Junctions. Transport and Road Research Laboratory (TRRL), SR-582, Crow throne, Berkshire, England (1980)
4. Glue, A.W.: Vereinfachtes Verfahren Zur Berechung signalgeregelter Knotenpunkte. Forschung Strassenbau und Strassenverkehrstechnik, Bonn **136**, 36–45 (1972)
5. Fellendorf, M., Vortisch, P.: Validation of the microscopic traffic flow model VISSIM in different real-world situations. In: 81st Annual Meeting of the Transportation Research Board, vol. 132, pp. 36–41, Washington D.C (2001)
6. Park, B., Schneeberger, J.D.: Microscopic simulation model calibration and validation: case study of VISSIM simulation model for a coordinated actuated signal system. Transp. Res. Rec. **2531**, 185–192 (1856)
7. Leyn, U., Vortisch, P.: Calibrating VISSIM for the German highway capacity manual. In: 94th Annual Meeting of the Transportation Research Board, vol. 1209, pp. 123–127, Washington D.C (2014)
8. Viti, F., Wolput, B., Tampere, C.M.J., Vandervelden, P.: Dynamic modelling of VISSIM's critical gap parameter at unsignalized intersections. Transp. Res. Rec. J. Transp. Res. Board **2395**, 12–20 (2015). Washington D.C.
9. Siddharth S.M.P., Ramadurai, G.: Calibration of VISSIM for Indian heterogeneous traffic conditions. In: 2nd CTRG, Procedia: Social and Behavioural Science, vol. 104, pp. 380–389 (2013)
10. Mathew, V.T., Radhakrishnan, P.: Calibration of micro-simulation models for non-lane-based heterogeneous traffic at signalized intersections. J. Urban Plann. Dev. **136**(1), 59–66 (2010)
11. Sangole, J.K., Patil, G.R.: Modeling vehicle group gap acceptance at uncontrolled T-intersections in Indian traffic. In: 93rd Annual Meeting of TRB, vol. 1789, pp. 156–162, Washington D.C (2014)
12. Rao, V.T., Rengaraju, V.R.: Modelling conflicts of heterogeneous traffic at urban uncontrolled intersections. J. Transp. Eng. **124**, 23–34 (1998)

Part VII
Real-time and Embedded Systems, Communication and Devices

Electromagnetic Band Structure Computation of Metamaterial/Air Composition from First Principle for Optical Filter Application

Bhaswati Das and Arpan Deyasi

Abstract In this paper, transfer matrix technique is used to compute the First Brillouin zone of DNG/air material composition for the application in photonic crystal. Three different types of physically realizable metamaterials are considered as the constituent of the periodic arrangement. Tuning of the Brillouin zone is made by suitable changing the structural parameters and coupling coefficient between forward and backward propagating waves for all three structures (paired nanorod, nano-fishnet with rectangular void, nano-fishnet with elliptical void). Results obtained from the first principle show the possibility of transition between perfect and quasi electromagnetic bandgap which is important for the possible application as photonic filter. Result is compared with the obtained band structure of conventional SiO_2/air composition.

Keywords Electromagnetic band structure · Brillouin zone · Dispersion relation · Metamaterial · Coupling condition · Structural parameters

1 Introduction

The dispersive effects can be seen when an electromagnetic wave propagates through a periodic structure along the direction of refractive index variation. This property holds for photonic crystal where permittivity of the dielectric material changes periodically. Thus, a novel property termed as electromagnetic/photonic bandgap [1, 2] is observed where some wavelengths of the incident spectra becomes restricted for propagation, and rest are allowed to pass [3]. Looking from a

B. Das (✉) · A. Deyasi
Department of Electronics and Communication Engineering,
RCC Institute of Information Technology, Kolkata, India
e-mail: bhaswati.das1991@gmail.com

A. Deyasi
e-mail: deyasi_arpan@yahoo.co.in

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_32

different angle, this structure may be recognized as optical band pass filter [4]. This physical manifestation of electromagnetic wave transmission can be explained by Bragg's principle of reflection [5]. This unique characteristic of photonic crystal is already utilized to fabricate photonic crystal fiber [6], and suitable choice of layer dimensions [7] gives an improve performance in communication comparing the performance presented by conventional optical fiber [8]. To construct optical transmitter [9], receiver [10], switch [11] as well as waveguide [12] photonic crystal is u sed. It is also used in quantum computation [13] and photonic integrated circuit [14].

Hillebrand introduces the bandgap in 2D photonic crystal structure by varying column roundness [15] using plane wave expansion method. To analyze the forbidden region of photonic crystal with different geometries finite-difference-time-do main method is used recently [16]. The width of the bandgap is calculated by Zhao [17] by using Bragg's reflection principle. To design large bandgap crystal evolutionary algorithm [18] and level-set method [19] are used. The innovation of semiconductor heterostructure gives uniqueness and provides increased performance as optical transmitting/emitting device [20]. This structure adds tunability of its filter characteristics [21].

Double negative materials are already used to reduce the return loss of some selected frequencies to improve performance o f ultra-wideband TEM horn [22]. Double negative medium is proposed for microwave applications by forming some complex artificial structure [23]. Novel binary diffractive lens is designed using DNG material for millimeter wave imaging system [24]. TE mode propagation inside multilayer planar waveguide at optical frequency is proposed [25] in recent past. Using fishnet structures with modified geometry, optical trapping is obtained at nanoscale [26] at near infrared wavelength. These results together suggest that DNG material is a known material in recent past for optical application. Hence its electromagnetic bandgap property needs investigation.

In this paper, electromagnetic bandgap is analytically computed for double negative material/air composition, and effect of grating lengths and coupling conditions are investigated. Result is compared with conventional system, and existence of complete and quasi photonic bandgaps are proved.

2 Mathematical Formulation

From wave equations, transfer matrix corresponding to the interface of the metamaterial and air can be obtained as

$$M_{1,2}^{T} = \frac{1}{t}\begin{pmatrix} 1 & r_{21,12} \\ r_{21,12} & 1 \end{pmatrix} \tag{1}$$

From the knowledge of wave-vector and layer dimensions, propagation matrix is given by

$$P_{1,2} = \begin{pmatrix} \exp[jk_{1,2}d_{1,2}] & 0 \\ 0 & -\exp[jk_{1,2}d_{1,2}] \end{pmatrix} \tag{2}$$

where *ki* is the wavevector in that layer, *di* is the propagation length in *i*th layer. Transfer matrix of the unit periodic block is given by

$$M = M_1^T P_1 M_2^T P_2 \tag{3}$$

For a perfectly periodic medium composed of N no. of blocks, total transfer matrix is given by

$$M_{tot} = M_N \tag{4}$$

Solution of the matrix gives eigenvectors which are actually the Bloch modes, hence we can write

$$\begin{bmatrix} a_n \\ b_n \end{bmatrix} = \exp(-jkz) \begin{bmatrix} a_{n-1} \\ b_{n-1} \end{bmatrix} \tag{5}$$

Using transfer matrix analysis

$$\begin{bmatrix} a_{n-1} \\ b_{n-1} \end{bmatrix} = M \begin{bmatrix} a_n \\ b_n \end{bmatrix} \tag{6}$$

A comparative study between Eqs. (5) and (6) gives the eigenvalue equation as

$$M \begin{bmatrix} a_n \\ b_n \end{bmatrix} = \exp(jk\Lambda) \begin{bmatrix} a_n \\ b_n \end{bmatrix} \tag{7}$$

where Λ is the periodicity of the grating. Considering exponential term as λ, we may write

$$\begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = 0 \tag{8}$$

for which solution is given by

$$\lambda = \frac{1}{2}(a+d) \pm \sqrt{\left[\frac{1}{2}(a+d)\right]^2 - 1} \tag{9}$$

Thus dispersion relation for the Bloch mode is

$$k = \frac{1}{\Lambda}\cos^{-1}\left[\frac{1}{2}(a+d)\right] \tag{10}$$

Equation (10) states that once the transfer matrix for the single unit block is determined, ω–k dispersion relation for the infinite periodic structure can easily be computed.

3 Results and Discussions

Using Eq. (10), equivalent Brillouin zone is computed and plotted for the DNG/air structure. Figure 1 depicts the band structure for different grating lengths (for 30, 45, and 60 μm) assuming the presence of coupling between forward and backward propagating waves. Result is plotted for paired nanorod (r • i = −0.3), nano-fishnet with rectangular void (r • i = −1), nano-fishnet with elliptical void (r • i = −4). From the plot, it is evident that quasi bandgap is obtained for rectangular void structure, but complete bandgap is obtained for the other two. For all the three cases, it is found out that the 'L' and 'G' points are separated maximum for higher grating length. Again, maximum number of optical forbidden zones is occurred for elliptical void structure, which ensures narrow bandpass filtering.

Figure 2 exhibits the profile for different coupling coefficients. We consider all three coupling conditions (low, medium, and high) and paired nanorod (−0.3 μm)/air composition for the first plot and observe a discontinuous band structure (Fig. 2a). Again for Fig. 2b, we get continuous band structure where we consider rectangular void structure. For elliptical void/air composition, discontinuity arises as seen in Fig. 2c. It may be noted in this context that the effect of coupling

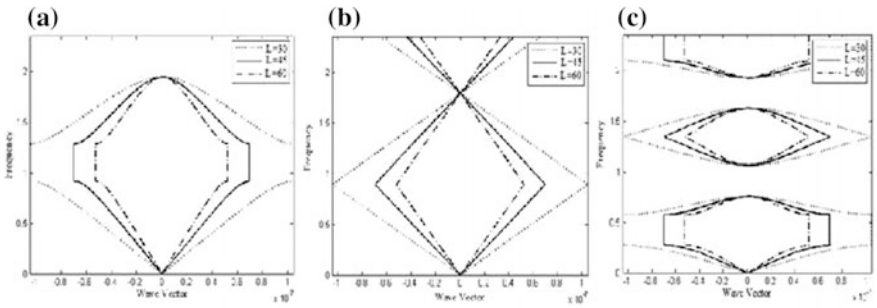


Fig. 1 Electromagnetic band structure for different metamaterial structure/air composition (**a** paired nanorod; **b** nano-fishnet with rectangular void; **c** nano-fishnet with elliptical void) with different grating lengths

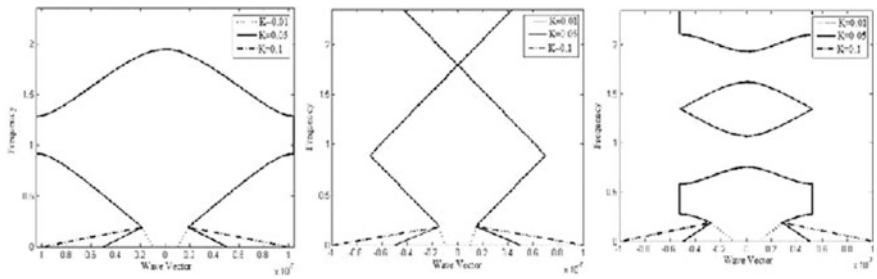


Fig. 2 Electromagnetic band structure for different coupling conditions with different metamaterials at different coupling conditions

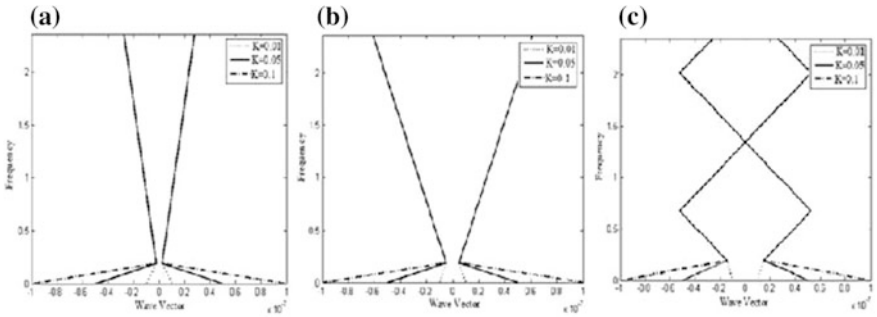


Fig. 3 Electromagnetic band structure for different metamaterial structure/air composition with different coupling conditions and different grating lengths (grating length is varied by changing the dimension of metamaterial)

condition is only significant at very lower frequency range. Lower coupling coefficient makes the origin of electromagnetic band much closer to the central point of wave-vector (k = 0), whereas higher magnitude of coefficient makes it further apart.

Figure 3 shows the result for the three different metamaterial structures with different grating lengths. This is made by changing the dimension of metamaterial (d1). Interestingly, it is seen fro m the plot that with change of grating length, complete bandgap can be tuned to incomplete one. Thus optical restriction of the desired frequency spectrum depends on the dimension of the structure.

Figure 4 represents the comparative study with the conventional SiO_2/air composition. By keeping the coupling coefficient and grating length constant, simulation is performed at different grating lengths independently. It may be noted that for restricting wave propagation, conventional system is better than rectangular void structure, but the spectrum zone is much wider in case of paired nanorod or elliptical void structure.

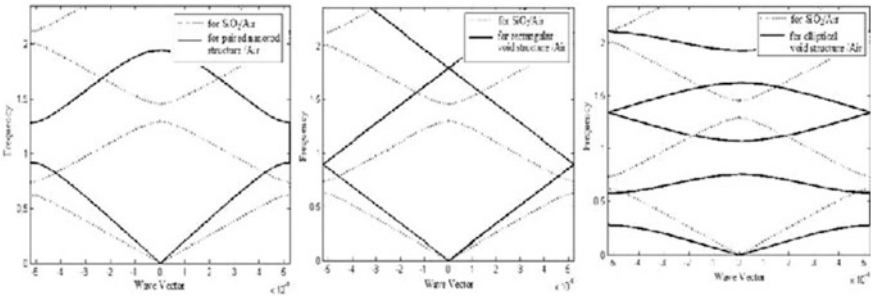


Fig. 4 Comparative study of electromagnetic band structure with conventional SiO_2/air composition

4 Conclusion

Electromagnetic band structure is calculated from dispersion relation using first principle, and result is plotted for first Brillouin zone. This reflects the physics of bandgap formation for the DNG material, and its possible utilization as 1D PhC. From the results we can see that both quasi band structure and complete bandgap formation are possible. Result reveals that the property can be effective for design of optical filter with desired restriction of certain propagating wavelength.

References

1. Yablonovitch, E.: Inhibited spontaneous emission in solid-state physics and electronics. Phys. Rev. Lett. **58**, 2059–2061 (1987)
2. Loudon, R.: The propagation of electromagnetic energy through an absorbing dielectric. J. Phys. A **3**, 233–245 (1970)
3. D'Orazio, A., De Palo, V., De Sario, M., Petruzzelli, V., Pru-denzano, F.: Finite difference time domain modeling of light amplification in active photonic bandgap structures. Prog. Electromagn. Res. **39**, 299–339 (2003)
4. Mao, D., Ouyang, Z., Wang, J.C.: A photonic-crystal polarizer integrated with the functions of narrow bandpass and narrow transmission-angle filtering. Appl. Phys. B **90**, 127–131 (2008)
5. Ozbay, E., Guven, K., Aydin, K: Physics and applications of photonic nanocrystals. Int. J. Nanotechnol. **1**, 379–398 (2004)
6. Russell, P.S.J.: Photonic-crystal fibers. J. Lightwave Technol. **24**, 4729– 4749 (2006)
7. Limpert, J., Schreiber, T., Nolte, S., Zellmer, H., Tunnermann, T., Iliew, R., Lederer, F., Broeng, J., Vienne, G., Petersson, A., Jakobsen, C.: High-power air-clad large-mode-area photonic crystal fiber laser. Opt. Express **11**, 818–823 (2003)
8. Hansryd, J., Andrekson, P.A., Westlund, M., Li, J., Hedekvist, P.O.: Fiber-based optical parametric amplifiers and their applications. IEEE J. Sel. Top. Quantum Electron. **8**, 506–520 (2002)
9. Szczepanski, P.: Semiclassical theory of multimode operation of a distributed feedback laser. IEEE J. Quantum Electron. **24**, 1248–1257 (1988)

10. Kalchmair, S., Detz, H., Cole, G.D., Andrews, A.M., Klang, P., Nobile, M., Gansch, R., Ostermair, C., Schrenk, W., Strasser, G.: Photonic crystal slab quantum well infrared photodetector. Appl. Phys. Lett. **98**, 011105 (2011)
11. Chen, J.C., Haus, H.A., Fan, S., Villeneuve, P.R., Joannopoulos, J.D.: Optical filters from photonic band gap air bridges. J. Lightwave Technol. **14**, 2575–2580 (1996)
12. Mekis, A., Chen, J.C., Kurland, I., Fan, S., Villeneuve, P.R., Joannopoulos, J.D.: High transmission through sharp bends in photonic crystal waveguides. Phys. Rev. Lett. **77**, 3787–3790 (1996)
13. Azuma, H.: Quantum computation with Kerr-nonlinear photonic crystals, J. Phys. D: Appl. Phys. **41**, 025102 (2008)
14. Bayat, K., Rafi, G.Z., Shaker, G.S.A., Ranjkesh, N., Chaud-huri, S.K., Safavi-Naeini, S.: Photonic-crystal based polarization converter for terahertz integrated circuit. IEEE Trans. Microw. Theory Tech. **58**, 1976–1984 (2010)
15. Hillebrand, R., Hergert, W., Harm, W.: Theoretical band gap studies of two-dimensional photonic crystals with varying column roundness. Physica Status Solidi (B) **217**, 981–989 (2000)
16. Popescu, D.G., Sterian, P.: FDTD analysis of photonic crystals with square and hexagonal symmetry. J. Adv. Res. Phys. **2**, 021105 (2011)
17. Zhao, J., Li, X., Zhong, L., Chen, G.: Calculation of photonic bandgap of one dimensional photonic crystal. J. Phys.: Conf. Ser. (Dielectrics 2009: Measurement Analysis and Applications, 40th Anniversary Meeting), **183**, 012018 (2009)
18. Preblea, S., Lipson, M.: Two-dimensional photonic crystals designed by evolutionary algorithms. Appl. Phys. Lett. **86**, 061111 (2005)
19. Kao, C.Y., Osher, S., Yablonovitch, E.: Maximizing band gaps in two-dimensional photonic crystals by using level set methods. Appl. Phys. B. **81**, 235–244 (2005)
20. Yin, Y., Huang, J.: High-speed visible light communication using light-emitting diodes embedded with photonic crystals. Progress in electromagnetic research symposium (2016)
21. D'souza, N.M., Mathew, V.: Tunable filter using ferroelectric-dielectric periodic multilayer. Appl. Opt. **54**, 2187–2192 (2015)
22. Weily, A.R., Esselle, K.P., Bird, T.S., Sanders, B.C.: Linear array of woodpile EBG sectoral horn antennas. IEEE Trans. Antennas. Propag. **54**, 2263–2274 (2006)
23. Foteinopoulou, S.: Photonic crystals as metamaterials. Phys. B.: Condens. Matter. **407**, 4056–4061 (2012)
24. Wang, Z.X., You, L.Z.: Design and analysis of double negative binary diffractive lens. International conference on microwave and millimeter wave technology (2010)
25. Shyroki, D.M., Lavrinenko, A.V.: Dielectric multilayer waveguides for TE and TM mode matching. J. Opt. A.: Pure Appl. Opt. **5**, 192–198 (2003)
26. Cao, T., Cryan, M.J.: Modeling of optical trapping using double negative index fishnet metamaterials. Prog. Electromagnet. Res. **129**, 33–49 (2012)

Mobility Prediction for Dynamic Location Area in Cellular Network Using Hidden Markov Model

Nilesh B. Prajapati and D.R. Kathiriya

Abstract To provide good quality of services to Mobile Users (MU) is main aim of every Cellular network. Radio bandwidth is critical resources which should be used optimally. More bandwidth is consumed due to frequent Location Update and paging. So, if we know the location of Mobile users in advance Location update and more paging can be reduced. In this paper, we implemented HMM method to predict current location of mobile user in Cellular network based on their previous mobility pattern and behavior. By the implementation, results shows that based on previous state of mobile user we can able to predict its current location. If previous state history is more than chance of getting accurate location of mobile user is higher.

Keywords Mobile users · Dynamic Location Area (DLA) · Cellular Network · Hidden Markov Model (HMM) · Location Update (LU) · Paging

1 Introduction

In current scenario, mobile users are increasing day by day as well as mobile technologies are also changing rapidly. As mobile users increase drastically in cellular network which increases signaling. Signaling requires too much bandwidth for all operation; like database access, LU, and Paging. To decrease signaling operations it is required to do better Mobile Network planning. Proper cellular network design and architecture are required for better Quality of Services (QoS). In GSM, for betterment different methods are available using which we can divide

N.B. Prajapati (✉)
Computer IT Engineering, B.V.M. Engineering College,
Gujarat Technological University, V.V. Nagar, Anand, India
e-mail: nbp_it53@yahoo.com

D.R. Kathiriya
Computer Center, Anand Agriculture University,
Anand 388110, Gujarat, India
e-mail: drkathiriya@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_33

the cellular networks into Location Area (LA). The size and shape of LA is more important in Cellular network. In fact, the size of an LA can be optimized to create a balance between the Location update rate and the expected paging rate within an LA. In other words, proper planning of LA reduces total Location management cost [1, 2]. There are two types of LA planning methods: Static and Dynamic.

For static LA planning methods two approaches are available never update [3] and Always Update [3]. Likewise Select-update approach is useful for DLA planning methods. In Select-update [3], the LU occurs only when certain predefine condition or constraints are violated. DLA methods are classified into State-based and Profile-based methods. By these methods we can able to locate Mobile User (MU) in the network whenever required by the network. Therefore, the location update/registration cost is reduced and the paging cost is also reduced up to measurable amount. Dynamic LA planning contains numbers of methods: Time-based [3–5], Movement-based [3–5], Distance-based [3–5], Cartesian product-based [4], User Profile-based methods [5, 6], Artificial Neural Network [6, 7], Simulated Annealing [8], Directed Graph [3], Data mining [9],etc. All these methods are useful for creation of dynamic LA for mobile user with some pros and cons. DLA is useful for getting location of users at predefined time. In this paper, DLA is used to find next possible cell movement of MUs by Hidden Markov Model (HMM) using their User Profile information. User profile information contains user's path, number of cells visited, and residential time within cells. If mobility prediction, next movement, of MUs are known in advance then we can provide better services, resource reservation, to MUs and reduce the call dropping rate. Mobility Prediction is also useful for cellular network for bandwidth management and call management.

The rest of paper is organized as follow. In Sect. 2, we describe introduction of Hidden Markov model. In Sect. 3, we describe implementation of HMM for mobility prediction. In Sect. 4, represent the results obtain using HMM, which shows how HMM used for mobility prediction. And at last in Sect. 5, we conclude the work with future enhancement.

2 Hidden Markov Model

In mobile network, number of users use the same path to reach their workplace or study place from their home or residential place. These kinds of users have fix movement pattern which we can utilize to reduce signaling and ultimately bandwidth. Using MUs mobility in nearest cells or in regular cells we can calculate their transition probabilities. As we have mobility history and behavior HMM can be used to predict current location of MUs in mobile network.

A Hidden Markov Model has two random processes. The first process is a Markov chain that is described by states and transition probabilities within the given network. The second process produces emissions observable at each moment, depending on a state-dependent probability distribution [10]. So, we have used HMM to obtain objective of the research using campus wireless trace dataset [11].

From the dataset, we can find out current Access Point (AP), last AP, and APs used by the MUs in their daily movement. Using mobility history, it is easy to calculate transition probabilities of every MUs. By the transition probability of MU, system can predict transition of MU from one AP to some other nearest AP. Sometime MU take several paths to reach their destination. But these scenarios happen in rare case like, traffic problem, etc. Transition probabilities of such a path are different than regular path.

Each Hidden Markov Model has five key elements: States, Observation Sequence, State transition probabilities, Observation State probabilities, and initial state probabilities which are useful to define an HMM completely [10] (Fig. 1).

Above figure shows that five elements, State (S), Observation sequence (O), initial state probability (π), state transition probability (A), and observation state transition probability (B).

In the System,

- $\{S_1, S_2, \ldots, S_N\}$ are the N State of any given system.
- $\{O_1, O_2, \ldots, O_N\}$ are the value of the observation sequence of the system.
- $\{\pi\}$ is the initial state probabilities so $\{\pi_i\}$ shows the probability of starting in state i.
- $\{a_{ij}\}$ are the state transition probabilities where a_{ij} denotes the probability of moving from state i to j.
- $\{b_{ik}\}$ are the observation state probabilities where b_{ik} is the probability of emitting symbol k at state i.

The notation $\lambda = (A, B, \pi)$ is often used to denote a HMM, that means discrete probability distributions of MUs in the cellular network. Using these parameters, if system wishes to predict location of any MU within the cellular network then HMM is built for that. But it is possible when mobility history data of MU's within the system.

In implementation, APs are considered as states of the system. From the MUs movement in each cell from current cells transition probabilities a_{ij} are calculated. Probable movements from current AP to next APs are considered as the observation sequences.

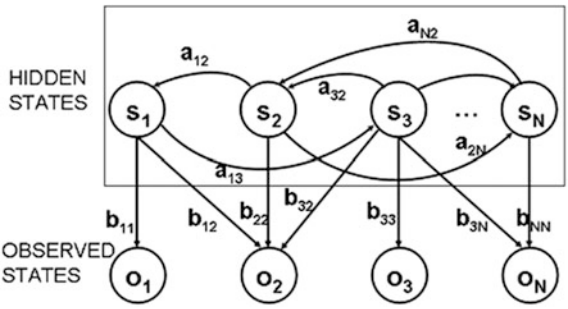


Fig. 1 A Hidden Markov Model [10]

3 Implementation

MUs mobility history is used for finding the next possible movement in cellular network. We have used a campus wireless trace dataset of Dartmouth university [11] for train the HMM network as well as a sample data to find location of users. This dataset contains wireless user histories such as user associations with APs, duration of association, etc. This dataset is very useful to get information such as the number of unique users, their association with different APs, the probability of transitions from one AP to another, etc.

There are several steps which are required to be followed to predict MUs location in mobile network. First, train the HMM using dataset based on how many previous states required observing to find the current location of MUs in mobile network. Based on that transition, probabilities are calculated using Eq. 1, which give probabilities of MUs frequent move in nearest cells. After that, for testing the train network sample data are taken from the dataset. Using the train network and transition probabilities we can get the true or false results for correctly identifying location of MUs in the network.

$$P(S_{t+1}|S_0, S_1, S_2, \ldots, S_t) = P(S_{t+1}|S_t) \text{ where } S_t \text{ is the state at time t} \quad (1)$$

From the dataset, seven user′s information are used in this implementation. We have trained the network up to previous seven states, for observing current location of MUs, and calculated transition probabilities for that. Our aim is to find out optimal state sequence and probability of an observed sequence. We have calculated percentage of true and false results of seven users for first to seven previous states using Eq. 2.

$$P(S_0, S_1, S_2, \ldots, S_T) = P(S_0) \prod_{(t=1)}^{T} P(S_t|S_{t-1})P(O_t|S_t) \quad (2)$$

4 Result

Below Fig. 2 shows the results of true and false prediction of MUs location of seven MUs in percentage. X-axis contains the previous states for which we calculate transition probabilities while Y-axis denotes the % of true and false prediction of users' future prediction. From the figure we can say that in majority case, approximately 80% of time, MUs future location can be predicted. In some cases results of the true prediction of users' locations are nearer to 90%.

Figure 3 shows the results of all seven MUs for prediction based on previous seven states. From that we can clearly say that if previous state information is more than possibility of MUs prediction in Mobile network is higher. X-axis contains the MUs whom future location required to predict while Y-axis denotes the % of true and false prediction of users' future prediction.

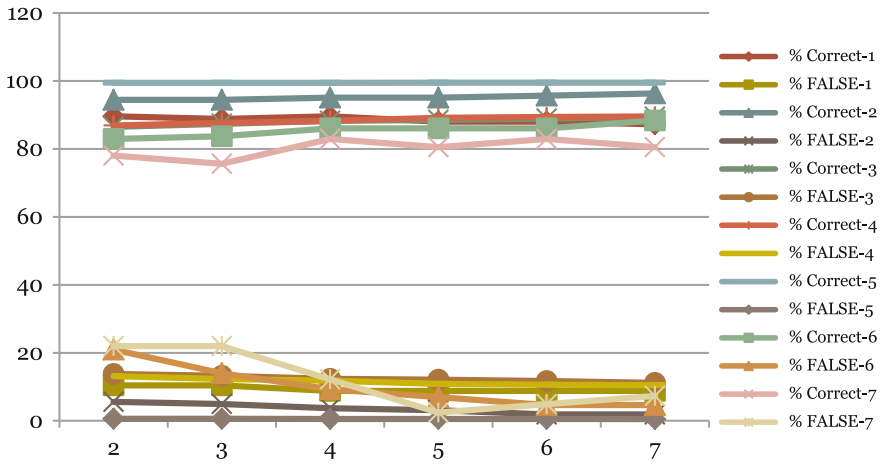


Fig. 2 % of true and false prediction of MUs based on previous states

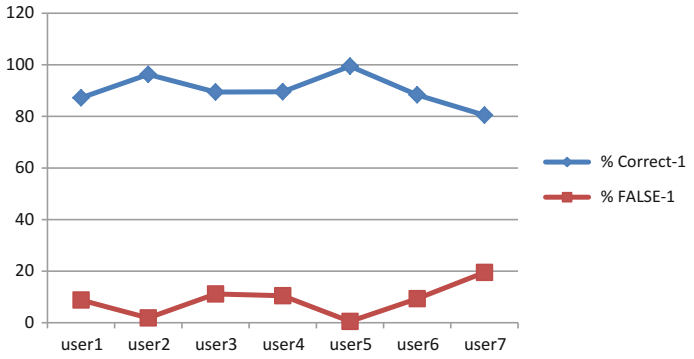


Fig. 3 % of true and false prediction of MUs based on previous seven states

Results of HMM gives information that, MUs previous state information (mobility behavior) gives more impact to find mobility patterns and future location in the network.

5 Conclusion and Future Work

Mobility prediction of any MU is possible using the Hidden Markov Model using MU's mobility history and behavior. In HMM methods previous state information play vital role to calculate transition probabilities so MUs history and same mobility pattern gives best results. In other words, for predicted and estimated users HMM gives good results.

In future, we can also check applicability of other similar regression algorithms and compare results. Also we can create generalize cluster irrespective of MUs, based on similar mobility pattern, to provide better services and reduce total Location Management cost, resource consumption of Cellular Network.

References

1. Mudaliar, K., Swadas, P., Prajapati, N.: Location management for cellular network using ant colony optimization. In: International Conference on Recent Trends on Computer Technology in Academia, 21–23 Apr 2012
2. Mudaliar, K., Prajapati, N.: User based personalization Scheme for dynamic location management. In: International Conference "ICNICT 11" at Karpagam University, Coimbatore, 15–16 Dec 2011
3. Lee, G., Chen, A.L.P.: The design of location regions using user movement behaviors in PCS systems. Multimedia Tools Appl. **15**(2), 187–202 (2001)
4. Markande, S.D., Bodhe, S.K.: Cartesian coordinate system based dynamic location management scheme. Int. J. Electron. Eng. Res. **2** (2009)
5. Sricharan, M.S., Vaidehi, V.: A dynamic distance based location management strategy utilizing user profiles for next generation wireless networks. In: First International Conference on Industrial and Information Systems, ICIIS2006, 8–11 Aug 2006, Sri Lanka
6. Vijay Kumar, B.P., Venkataram, P.: Prediction-based location management using multilayer neural networks. Indian Inst. Sci. **82**, 7–21 (2002). © Indian Institute of Science
7. Singh, J.A.P., Karnan, M.: A dynamic location management scheme for wireless networks using cascaded correlation neural network. Int. J. Comput. Theory Eng. **2** (2010)
8. Zheng, J., Regentova, E., Srimani, P.K.: Dynamic location management with personalized location area of future PCS network. In: 6th International Workshop on Distributed Computing IWDC 2004, India
9. Aoudjit, R., Belkadi, M., Daoui, M., Chamek, L.: Mobility prediction based on data mining. Int. J. Database Theory Appl. **6**(2) (2013)
10. Kouemou, G.L.: History and Theoretical Basics of Hidden Markov Models. http://www.intechopen.com
11. Crawdad: Wireless Traces from Dartmouth. http://crawdad.cs.dratmouth.edu/
12. Rong, C., Senmiao, Y.: Distributed and Dynamic Location Area for PCS. IEEE, 1-4244-0463-0/06
13. Xie, H., Tabbane, S., Goodman, D.J.: Dynamic location area management and performance analysis. In: Proceeding of the 43rd IEEE Vehicular Technology Conference
14. Landfeldt, B., Kolodziej, N.: A dynamic location management scheme based on individual metrics and coordinates. In: Proceedings of IEEE WITSP04, Adelaide, Australia, Dec 2004
15. Scourias, J.: Dynamic Location Management and Activity-Based Mobility Modelling for Cellular Networks. Waterloo, Ontario, Canada, 1997 © John Scourias (1997)
16. Foughali, L., Talbi, E.-G.: A Parallel Insular Model for Location Area Planning in Mobile Networks. IEEE, 978-1-4244-1694-3 (2008)
17. Pierre, S., Houeto, F.: Assigning cells to switches in cellular mobile network using taboo search. IEEE Trans. Syst. **32**(3), 351–356 (2002)
18. Akyildiz, I.F., McNair, J., Ho, J., Wang, W.: Mobility Management in Next Generation Wireless Systems. Broadband and Wireless Networking Laboratory School of Electrical and Computer Engineering Georgia Institute of Technology, Atlanta
19. Munguia-Marcario, M., Munoz-Rodriguez, D., Molina, C.: Optimal adaptive location area design and inactive location area. In: Proceedings of 47th IEEE Vehicular Technology Conference, vol. 1, pp. 510–514 (1997)

20. lBejerano, Y., Smith, M.A., (Seffi) Naor, J., Immorlica, N.: Efficient location area planning for personal communication systems. IEEE/ACM Trans. Netw. **14**(2), (2006)
21. Wan, G., Lin, E.: Cost reduction in location management using semi-realtime movement information. Baltzer J. (1998)
22. Scourias, J., Kunz, T.: A Dynamic Individualized Location Management Algorithm. Department of Computer Science, University of Waterloo. 0-7803-3871-5/97/$10.00 © 1997 IEEE
23. Giner, V.C.: State of the art in Location Management procedures. Information Society Technologies (IST)—6th Framework Programme

Neighbor Constraint Traffic Centric Distributed Sinkhole Detection and Mitigation Approach for Quality of Service Improvement in Wireless Sensor Networks

K. Devibala, S. Balamurali, A. Ayyasamy and M. Archana

Abstract In Wireless Sensor Networks (WSN), the problems of sinkhole detection and mitigation have been studied through various methods, but suffer with the problems of high complex detection and overhead introduced by detection process. The Quality of Service (QoS) has been degraded greatly by the impact of sinkhole attack because the malicious node can read much information about the source or the communications; this helps malicious nodes to perform various kinds of attacks. In order to overcome such drawbacks of network security, in this paper propose a Neighbor Constraint Traffic Centric (NCTC) Method for sinkhole detection to support QoS development of WSN. The problem of multiple identifications is resolved by verifying the transmission performed through other neighbors of the malicious node. From the traffic incurred in the neighbor of malicious node is used to verify the location or traffic sent by the malicious node. The method produces efficient results in sinkhole detection and increases the throughput and also the packet delivery ratio and reduces the frequency of sinkhole detection.

K. Devibala (✉)
Department of Computer Science, Ayya Nadar Janaki Ammal College, Sivakasi, Tamil Nadu, India
e-mail: sreebalahoney@gmail.com

S. Balamurali
Department of Computer Application, Kalasalingam University, Virudhunagar, Tamil Nadu, India
e-mail: sbmurali@rediffmail.com

A. Ayyasamy · M. Archana
Faculty of Engineering and Technology, Department of Computer Science and Engineering, Annamalai University, Annamalai Nagar, Chidambaram, Tamil Nadu, India
e-mail: samy7771@yahoo.co.in

M. Archana
e-mail: archana.aucse@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_34

Keywords Wireless sensor networks · Sinkhole attacks · Neighbor-based detection · Traffic centric method · QoS

1 Introduction

The Wireless Sensor Network (WSN) [1] has the capability to be deployed in any point of time without any preparation. The WSN has a number of nodes which has mobility nature and could be used to form a network at many situations like war fields. The WSN are more useful in deploying the network in war filed and collect the information which will be used to perform various activities. Similarly the WSNs are useful in various situations to collect information. The nodes of the network perform both reception and transmission of packets for communication [2–4].

The co-operative communication is performed with the support of all the nodes of the network and there are some nodes which try to capture the packets being forwarded and learnt something from the packets [5, 6]. The learned information can be used to perform various network threats and the malicious nodes are capable of sending the packet direct to the base station. This creates a set of energy holes around the base station which spoils the network life time and increases direct traffic to the base station [7, 8].

The rest of the paper is classified as follows, with Sect. 2, which introduces the set of core requirements for WSN test bets which help to understand the pragmatic deployment of WSN design and survey approach. Section 3 elaborates on the Proposed Architecture and based on the three algorithms clearly explained. Section 4 the NCTC is evaluated over WSN sinkhole detection and mitigations and its performance evaluated. Finally, concluded the paper in Sect. 5.

2 Literature Survey

There are many approaches those have been discussed in literature and we present some of them here which are relevant to the problem statement. Detection of wormhole attacks in WSN using range-free localization [9, 10] suggests two wormhole detection procedures for WSNs, based on concepts working in a kind of range-free localization methods: one of the approaches performs the detection at the same time with the localization procedure, and the other operates after the conclusion of the location discovery protocol. Both approaches are effective in detecting wormhole attacks, but their presentation is fairly sensitive to shadowing effects present in the radio channels.

Detection and correction of sinkhole attack with novel method in WSN using NS2 Tool [11] uses a sequence number-based sinkhole detection approach [12]. The dispatcher node first requests the sequence number with the route request

(RREQ), if the node receives this RREQ then replies its sequence number with route reply (RREP) message. Broadcasting node checks whether this will match with sequence number in its routing table. If it matches then information will be shared otherwise it will allocate the sequence number to the node. If the node recognizes the sequence number then the node will go through the network otherwise it will be eliminates from the network [13].

Detection of sinkhole attack in wireless sensor networks [14, 15] proposes a Sybil attack detection scheme which initially uses the consistency of data to find the group of suspected nodes. Then, the intruder is documented efficiently in the group by checking the network flow information. The proposed algorithm demonstrates and evaluated by using numerical analysis and simulations. Therefore, accuracy and ability of algorithm would be verified.

3 Proposed NCTC Method

The proposed NCTC method has the following stages namely: Neighbor Centric Data Collection, Traffic Inference, and Sinkhole Attack Detection. The neighbor centric data collection is performed to collect the neighbor details and the traffic inference is performed to compute the participation of other nodes in the network, the illustration of the proposed work is shown in Fig. 1. The sinkhole detection approach detections the presence of sinkhole and malicious nodes of the network.

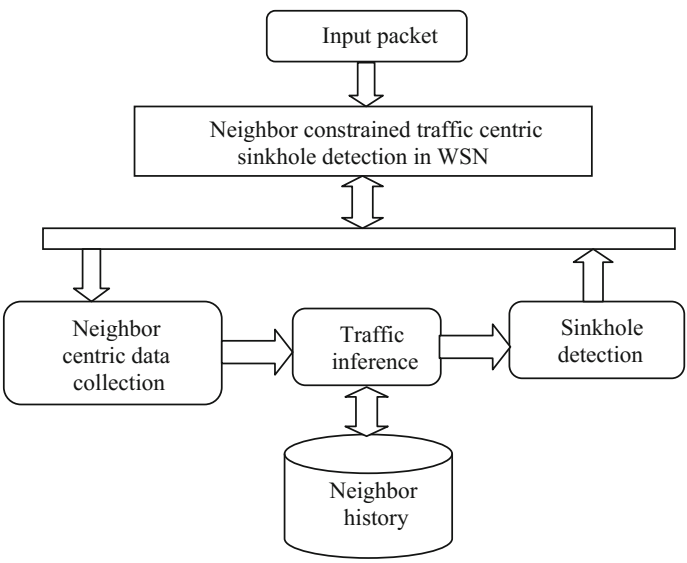


Fig. 1 Block diagram of proposed architecture

3.1 Neighbor Centric Data Collection

The node of the WSN performs the data collection whenever it has a packet to be transmitted. The node constructs a Data Collection Message (DCM), and broadcasts into the network. The nodes are located in the broadcast range and then source will accept the packet and reply with the Data Collection Message Respond (DCMR) message. The reply has various information about their location, transmission range, number of neighbors, and their information, number of transmission performed and received. The features from the message will be extracted and stored to perform sinkhole detection later as shown in Algorithm 1.

Algorithm 1

Input: NULL
Step 1: Start
Step 2: Construct DCM
Step 3: Broadcast DCM
Step 4: While (True)
Receive DCM-Reply Message.
Extract the features from DCM-Reply.
Number of neighbors TN = $\int \sum Node \in DCMR$
NodeID = $\int \sum Node.NodeID \in DCMR$
NLoC = $\int_{i=1}^{TN} Location(NodeID) \in DCMR$
TNR = $\int_{i=1}^{TN} TNR(NodeID) \in DCMR$
TNT = $\int_{i=1}^{TN} TNT(NodeID) \in DCMR$
ED = $\int_{i=1}^{TN} ED(NodeID) \in DCMR$
End
Step 5: Generate Node History
For all node id NID_i from NID
ND (i) = {TNN, NodeID, NLoC, TNR, TNT, ED}.
End.
Step 6: Stop.
Output: Neighbor Information NI

3.2 Traffic Inference

From the collected information about the neighbor nodes, we compute the traffic inference which represents the possibility of the neighbors participated in the transmission. There may be a malicious node located and could communicate

directly to the sink node or base station. In order to monitor such a flow we compute the traffic inference which could be computed from the transmission performed by other nodes of the network. The inference is computed using the number of transmission performed by a single node which is suspected by the sinkhole detection process.

Algorithm 2:
Input: Neighbor information NI
Step1: Start
Step2: For each neighbor of Node N_i
Compute Traffic Rate Inference TRI.

$$TRI = \int ND\left(\frac{(TNR+TNT)}{100}\right)$$

End.
Step3: Compute traffic rate of suspect node

$$TRS = \int ND\left(\frac{(TNR+TNT)}{100}\right)$$

Step4: if TRI<TRS then
Return true.
Else
Return false.
End.
Step5: Stop.
Output: Inference Flag IF.

3.3 Sinkhole Detection

The sinkhole detection is performed using the data collection and traffic inference model. The sinkhole detection is performed by each node of the network and the node identifies the malicious node as follows: Each node verifies the presence of multiple appearance of id in different nodes neighbor table and if there exist any id located in different nodes neighbor information then the traffic inference will be computed to conclude the sinkhole attack node. If there are many appearance of the same id in different nodes neighbor details the same is verified with the location details and inference model.

Algorithm 3:
Input: Neighbor information NI
Step1: Start
Step2: for each neighbor N_i from NI
Check for the appearance in multiple tables.
Count C= $\int_{i=1}^{size(NI)} Count(Ni, NI)$
If C> $\frac{2}{3} \times size(NI)$ then
Perform traffic inference TI
Flag F = TI (Ni, NI).
If F==True then
Declare Ni as malicious
Else
End.
End.
Step3: Stop
Output: NULL

4 Performance Evaluation

The proposed NCTC-based sinkhole detection and mitigation technique has been implemented in Network Simulator 2 (NS2) and tested for its effectiveness in all the measures of quality of service. The method has been simulated with different scenarios with different number of nodes. The simulations were agreed out using a WSN atmosphere consisting of 100 wireless nodes over a simulation area of 1000 × 1000 m flat space functioning for 60 s of simulation time. In the simulation each node has been considered with same set of transmission range of 100 m and the malicious nodes are assigned with the transmission range of 500 m. Table 1 illustrates the simulation details. The method has been simulated for its effectiveness in sinkhole detection with different topology and simulation parameters.

Table 2 shows the comparison of results produced by different methods in most important factors of quality of service by 100 nodes.

Table 1 The constraint used in simulation

Constraint	Value
Version	Network simulator-allinone 2.34
Protocol	NCTC
Area	1000 m × 1000 m
Transmission range	100 m
Traffic model	UDP, CBR
Packet size	512 bytes

Table 2 Shows the comparison results

S. No.	No. of nodes	Protocol	Detection rate		Throughput	PDF
			False +ve	False −ve		
1	100	Range free	7.6	6.7	79.0	8.4
2	100	SeRA	5.4	4.8	84.0	7.3
3	100	Polygon based	4.6	4.3	90.0	6.0
4	100	G-Hazard	3.5	2.5	92.0	6.9
5	100	NCTC	0.4	0.3	99.6	1.2

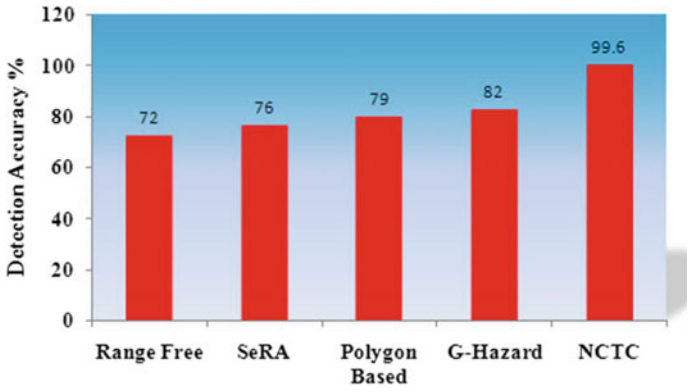


Fig. 2 Comparison of sinkhole detection accuracy

Figure 2 shows the comparison of sinkhole detection accuracy. The result reveals that the proposed method has produced more efficient detection accuracy than the other approaches.

4.1 Throughput Performance

Throughput is measuring the rate of success of data packets received at the destination. The throughput is measured in terms of number of bits per second (bps) or number of data packets per second (dps). Figure 3 shows the overall throughput ratio of different methods and it is clear that the proposed NDE method has achieved higher throughput than other methods.

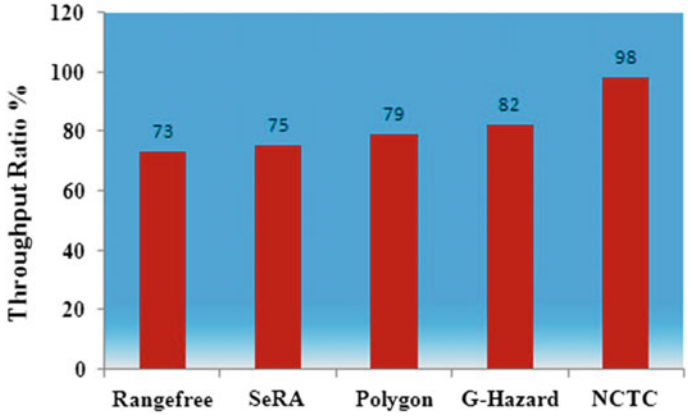


Fig. 3 Throughput ratio of different methods

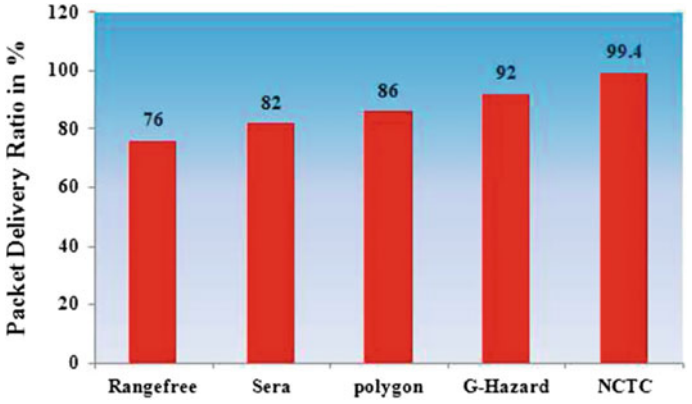


Fig. 4 Packet delivery ratio

4.2 Packet Delivery Ratio

Packet Delivery Ratio (PDR) is the standard metrics for measuring the performance of proposed protocol. The PDR is the rate of measures, the number of data packets received at a destination according to the number of packets produced in the source node. The average PDR is measured by

$$\text{PDF} = \text{Number of packets arrived} / \text{Number of packets dispatch} * 100.$$

Figure 4 shows the performance of packet delivery ratio of different algorithms and it indicates that the proposed NCTC method has higher packet delivery ratio than the other methods.

5 Conclusions

In this paper, a NCTC method for sinkhole detection approach is proposed to get better quality of service in WSN. The node collects the information about the neighbors and their location, number of transmissions, and receptions engaged. Similarly compute the traffic rate inference based on the two measures the node is decided as genuine or malicious. The detection process is executed based on the traffic introduction factor which is calculated based on packets being sent and received. Using the NCTC a single node routes to reach the sink node or base station is computed. From the available routes, the routes are verified to conclude the node is being malicious or not. The proposed method has reduced the overhead generated by distributed sinkhole detection process and produced efficient results.

References

1. Yu, F., Park, S., Tian, Y., Jin, M., Kim, S.: Efficient hole detour scheme for geographic routing in wireless sensor networks. In: 2008 IEEE International Conference on Vehicular Technology Conference (VTC), pp. 153–157 (2008)
2. Ganapathy, S., Kulothungan, K., Muthurajkumar, S., Vijayalakshmi, M., Yogesh, P., Kannan, A.: Intelligent feature selection and classification techniques for intrusion detection in networks: a survey. EURASIP J. Wirel. Commun. Netw. **23**(1), 1–16 (2013)
3. Vijayakumar, P., Azees, M., Kannan, A., Jegatha Deborah, L.: Dual authentication and key management techniques for secure data transmission in vehicular ad hoc networks. IEEE Trans. Intell. Transp. Syst. **17**(4), 1015–1028, (2015)
4. Ayyasamy, A., Venkatachalapathy, K.: Context aware adaptive fuzzy based QoS routing scheme for streaming services over MANETs. Wireless Netw. **21**(2), 421–430 (2015)
5. Shin, I., Pham, N.D., Choo, H.: Virtual convex polygon based hole boundary detection and time delay based hole detour scheme in WSNs. In: Human Interface and the Management of Information. Designing Information Environments, pp. 619–627. Springer Publication (2009)
6. García-Otero, M., Población-Hernández, A.: Secure neighbor discovery in wireless sensor networks using range-free localization techniques. Int. J. Distrib. Sensor Netw. (2012)
7. Choi, M., Choo, H.: Bypassing hole scheme using observer packets for geographic routing in WSNs. In: International Conference on Information Networking (ICOIN), pp. 435–440. IEEE (2011)
8. Sheela, D., Naveen, K.C., Mahadevan, G.: A non cryptographic method of sink holeattack detection in wireless sensor networks. In: International Conference on Recent Trends in Information Technology (ICRTIT), pp. 527–532. IEEE (2011)
9. Teng, L., Zhang, Y.: SeRA: a secure routing algorithm against sink-hole attacks for mobile wireless sensor networks. In: Second International Conference on Computer Modeling and Simulation, pp. 79–82. IEEE (2012)
10. García-Otero, M., Población-Hernández, A.: Detection of wormhole attacks in wireless sensor networks using range-free localization. In: 17th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), pp. 21–25. IEEE (2012)
11. Singh, T., Arora, H.K.: Detection and correction of sink-hole attack with novel method in WSN using NS2 tool. Editorial Preface, Citeseer **4**(2) (2013)
12. Singh, H., Malik, G.: Approaches to wireless sensor network: security protocols. World Comput. Sci. Inf. Technol. J. (WCSIT) **1**(7), 302–306 (2011)

13. Shukla, P.K., Silakari, S., Bhadoriya, S.S.: Network security scheme for wireless sensor networks using efficient CSMA MAC layer protocol. In: Sixth International Conference on Information Technology: New Generations, pp. 1579–1580. IEEE (2009)
14. Krontiris, I., Giannetsos, T., Dimitriou, T.: Launching a sinkhole attack in wireless sensor networks; the intruder side. In: International Conference on Wireless and Mobile Computing Networking and Communications, pp. 526–531. IEEE (2008)
15. Chen, C., Song, M., Hsieh, G.: Intrusion detection of sinkhole attacks in large-scale wireless sensor networks. In: International Conference on Wireless Communications Networking and Information Security (WCNIS), pp. 711–716. IEEE (2010)

Moving Object Detection Using Local Binary Pattern and Gaussian Background Model

A.P. Athira, Midhula Vijayan and R. Mohan

Abstract It has been several years Background subtraction techniques were put into use in vision and image applications for motion detection. However, most of the methods fall short of providing fine results due to dynamic backgrounds, illumination variation, noise, etc. Uniqueness of the proposal is construction of a steady background from a video sequel. In the editorial, proposal is to develop a steady background representation from a certain video sequel. The background is updated on arrival of each frame. For detecting moving objects, the constructed background has been compared with diverse frames of the video sequel. For this, the background model is developed using combination of Local Binary Pattern (LBP) and Gaussian averaging. Gaussian averaging employs different forms that occur with time to confines the underlying opulence of the background. Likewise, a spatial region of hold is used by LBP. The projected proposal depends on spatio-temporal forms occurring with time to fabricate a suitable model background. Efficacy of the projected proposal is established by comparing the outcomes with some of the existing avant-garde background subtraction methods on open standard records.

Keywords Gaussian model · Local Binary Pattern · Object detection

1 Introduction

Dynamic detection of object from video sequels [1] is a major difficulty in Image processing and computer vision. Accurate recognition of motion in objects is vital for several applications of artificial intelligence and image processing. Dynamic object

A.P. Athira (✉) · M. Vijayan · R. Mohan
Department of CSE, NIT Trichy, Trichy 620015, Tamil Nadu, India
e-mail: athiraap.19@gmail.com

M. Vijayan
e-mail: midhula91@gmail.com

R. Mohan
e-mail: rmohan@nitt.edu

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_35

detection is being used extensively in the areas like robotics, activity recognition [2], video surveillance [3], gait and face-based human recognition [4], etc.

Different object detection methods generally used are Background subtraction, frame differencing, and Optical flow. Out of these, Background subtraction has been the most accepted and popular method for detecting an object. Background subtraction constitutes of two basic steps-background modeling and subtraction [5]. Frame differencing technique for object detection uses a reference background image for comparing frames. Places, where changes are seen are identified as dynamic object in the case of frame differencing.

In the Background Subtraction (BGS) model proposed by Heikkila and Olli [6], both the concepts of frame differencing and background updation had been used. In this, background is updated according to current frame. Nevertheless this background model is rigid against non-static background, noise, and illumination variation. Based on this concern, Wren et al. [7] proposed a flexible BGS method by means of running Gaussian average. In this researchers have represented background autonomously at every pixel position by correcting a Gaussian pdf (probability distribution function) above a series of preceding frames. By means of intensity comparison of every pixel positions in the marked frame and the consequent factors of Gaussian pdf, the positions of the dynamic objects in a frame are attained.

So as to model a background correctly, a multi-valued background scheme is required. Stauffer and Grimson [8, 9] projected a scheme where multiple Gaussian pdfs are used to model pixels at a specific position of an image frame. This method gives improved results in non-static backgrounds and is found flexible and robust. Of late, this model was modified incorporating dynamic textures for detecting dynamic objects from non-static backgrounds [10]. From these, it can be concluded that the methods mentioned above do not consider spatio-contextual information for detecting changes and are pixel based. Such a spatio-contextual BGS scheme that uses non-altering textural property, LBP histogram has been scrutinized in [11].

From these analyses, it is obvious that various techniques for object detection applying background subtraction had been suggested by several analysts in the area of Image Processing. As in the case of everything, these algorithms have their own advantages and drawbacks. It can be brought to a conclusion that, for a good background subtraction technique to be designed, it requires construction and updation of a steady background which can exemplify the background of a video. It must also distinguish the effect of noise and dynamic background.

In this work that is proposed, efforts has been taken to merge the benefits of Gaussian model based background construction and the idea of Local binary pattern for non-static object detection from a video sequel. Therefore the model can be termed as a spatio-temporal model. Here, the proposal is a pioneer BGS method to detect non-static objects from static and dynamic backgrounds. Uniqueness of the model is in developing an unwavering spatio-temporal background scheme from a certain video sequel which proficiently illustrates the features of background in the sight. It is also noticed that the projected model offers better outcomes against small variations in light, noise, and dynamic background. The efficacy of the scheme proposed has been established by comparing with the BGS models using threshold,

and two different Gaussian mixture models. The observation is that this proposal offers better outcome in comparison with other existing methods considered.

As objects in a video are usually moving objects, this kind of variability changes are pretty evident. When an object shifts from one frame to another throughout the line of sight of a camera, image of an entity may change significantly. Listed below are the characteristic challenges of background subtraction are Dynamic Background, Occlusion, Clutter, Camouflage, Presence of Shadow, Motion of the Camera and Video Noise.

2 Proposed Algorithm

Here the proposal is a contemporary Background subtraction model that is capable of detecting dynamic objects in a video sequel by proficiently lessening the consequences of illumination difference, camera jitter, and moving backgrounds within the sequel. Here, in this projected model, the background is created by scrutinizing series of linearly dependent precedent image frames by means of Gaussian Mixture Model and Local binary pattern.

2.1 Local Binary Pattern

For texture matching in computer vision, one of the vital attribute descriptor is Local Binary Pattern (LBP). It is a potent way of texture description. LBP operator tags pixels of a scene by thresholding the zone (neighbors) of every pixel along with the middle value and by arraying the outcome to binary codes [12, 13]. Local Binary Pattern has few striking properties like gray scale invariance, non-parametric computational simplicity, illumination invariant and high discriminative.

The basic one has eight pixels as neighbors, which are easily extensive to contain a bigger neighborhood with any number of pixels. In LBP, a middle pixel is taken and its M neighbors are considered. Then the difference between middle pixel and its M neighbor pixels are computed to get the M-bit binary number. If the neighborhood value is less than central pixel, the value is 0 and otherwise 1. From this arose 2^{M} discrete decimal values for binary pattern. The process is as illustrated in the figure beneath. The value of the middle pixel is 7. Here we compares from the upper pixel and maintains a clockwise direction to cover all its neighbors.

$$\mathrm{LBP}_{\mathrm{M,R}}(c_x, c_y) = \sum_{q=0}^{M-1} \mathrm{T}\left(\mathrm{I_q} - \mathrm{I_c}\right) \tag{1}$$

$$T(\alpha) = \begin{cases} 1 & \alpha \geq 0 \\ 0 & \alpha < 1 \end{cases}$$

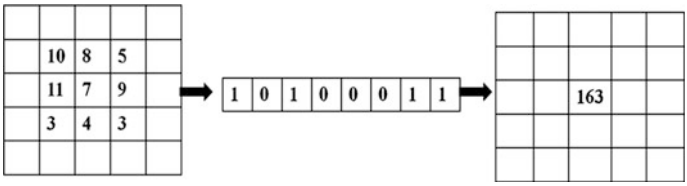


Fig. 1 Local binary pattern

Here I_c is the central pixel value with co-ordinate (c_x, c_y) and I_q is the neighboring value of the pixels in a circle of radius 1 unit (Fig. 1).

2.2 Background Construction

The primary step for any foreground segmentation method is background construction. The steadiness of the background model constructed decides accuracy of object detection. The accurateness of segmentation gets corrupted if the background constructed is degraded by noise. Even the dynamic background reduces the quality of segmented foreground. Hence, it is very essential to put up a vigorous background construction scheme for precise moving object detection.

In the projected proposal, the model background is constructed by taking into account the series of image frames that are previously available. The key concept of the proposed work is the dependent pixels of similar texture in the background at a certain position of the image that follows Gaussian distribution at temporal course. This scheme is constructed using intensity values of these dependant pixels. The dependency of pixels is checked using LBP histogram. For this reason, in the projected proposal we have utilized LBP histogram to verify whether the pixel in a certain position of the background model formed is reliant on the pixel in the equivalent position in the succeeding image frames. Initially we set the first frame as background.

$$\mu(\gamma, \delta) = \mathrm{I}_1(\gamma, \delta) \tag{2}$$

Here $\mu(\gamma, \delta)$ is the background model and $I_1(\gamma, \delta)$ is the first frame. We test the similarity between the pixels at location (γ, δ) using the following steps.

Step 1: Construct the LBP histogram for each pixel using binary masks for background image and target image. LBP histogram for each pixel is computed over neighboring pixels in a circular region.

Step 2: Find the similarity between target histogram and background histogram corresponding to the pixel value (γ, δ). The similarity between the histograms can be find out using the concept of correlation

$$D(H_t, H_b) = \frac{\sum_x (H_t(x) - \overline{H_t})(H_b(x) - \overline{H_b})}{\sqrt{\sum_x (H_t(x) - \overline{H_t})^2 \sum_x (H_b(x) - \overline{H_b})^2}} \tag{3}$$

Here $D(H_t, H_b)$ is similarity measure, H_t is the target LBP histogram corresponding to particular pixel (γ, δ), and H_b is the background LBP histogram corresponding to the same pixel (γ, δ).

Step 3: If the value of $D(H_t, H_b)$ does not deviate from 1, then it means the pixels are similar, and the corresponding pixel is taken as of backgrounds.

If a pixel in the position (γ, δ) is noticed alike with the constructed background then the model background is updated by taking into consideration the notion of Gaussian averaging. Background can be updated using the given formula,

$$\mu_t(\gamma, \delta) = \frac{(i-1)\left(\mu_{(t-1)}(\gamma, \delta)\right) + I_t(\gamma, \delta)}{i} \tag{4}$$

where i correspond to the number of dependant pixels already occurred at location (γ, δ) at time t, $I_t(\gamma, \delta)$ is intended frame and the background created is $\mu_{(t-1)}(\gamma, \delta)$. Background model is kept same, if it is observed to be independent.

$$\mu_t(\gamma, \delta) = \mu_{(t-1)}(\gamma, \delta) \tag{5}$$

and

$$var(\gamma, \delta) = \frac{1}{i} \sum_{k=0}^{i} |I_t(\gamma, \delta) - \mu_t(\gamma, \delta)|^2 \tag{6}$$

where $var(\gamma, \delta)$ represents the variance for pixel location (γ, δ).

2.3 Moving Object Detection and Background Updation

Updation of the background frame is necessary to acclimatize the reference image with the variations in the background scene. In order to detect a moving object precisely, it is essential to create a fresh background image all over again. As a result, the method of updating background is one of the main processes in dynamic object detection. This is performed using background subtraction.

It is necessary to take out the exact figure of dynamic objects in a scene. Here, for recognition of non-static objects firstly we take the background as initial frame. The pixel can be categorized as background or object, at any time t. Object can be represented as 1 and background can be represented as 0.

$$R_{t+1} = \begin{cases} 1 & if \quad abs\left(\frac{I(\gamma,\delta)-\mu_t(\gamma,\delta)}{\sqrt{var(\gamma,\delta)}}\right) \geq T_h \\ 0 & Otherwise \end{cases} \tag{7}$$

Here the absolute value is represented by abs(). At times the detected region is not connected so for better result we have used some morphological techniques. T_h represents a manually given threshold value. The concept of morphological closing on an image is used in the proposed work to remove small dark spots and connect small bright cracks. Updation of background is the next stage of background subtraction. In this proposal, the updation of the background model at each time (t + 1) has been performed as below,

$$\mu_{(t+1)}(\gamma,\delta) = \begin{cases} \eta I_{t+1}(\gamma,\delta) + (1-\eta)\mu_t(\gamma,\delta) & if\ R_{t+1}(\gamma,\delta) = 0 \\ \beta I_{t+1}(\gamma,\delta) + (1-\beta)\mu_t(\gamma,\delta) & Otherwise \end{cases} \tag{8}$$

and

$$var_{(t+1)}(\gamma,\delta) = \begin{cases} \eta(I_{t+1}(\gamma,\delta) - \mu_t(\gamma,\delta)^2 + (1-\eta)var_t(\gamma,\delta), & if\ R_{t+1}(\gamma,\delta) = 0 \\ \beta(I_{t+1}(\gamma,\delta) - \mu_t(\gamma,\delta)^2 + (1-\beta)var_t(\gamma,\delta) & Otherwise \end{cases} \tag{9}$$

where $var_t(\gamma, \delta)$ represents the variance of Gaussian model at pixel location (γ, δ). η represents the constant value that varies from 0 to 1 and changes with every frame, η is the ratio of total number of background pixels by total number of pixels. In order to adjust swiftly for the pixels identified as background and gradually for the pixels identified as foreground, β has to be much lesser than η and β usually is zero.

3 Experiments and Evaluation

The projected proposal is executed in python with OpenCV and is run on Intel Core i3 with Ubuntu OS and 4 GB RAM. For ascertaining efficacy of the projected method, trials have been carried out on four video sequels. The chosen video sequels are 'Water Surface', 'PETS2006', 'Traffic' and 'Curtain'. So as to legitimate the projected proposal, results achieved from the method is compared with two different Gaussian algorithms [14, 15] and thresholding method.

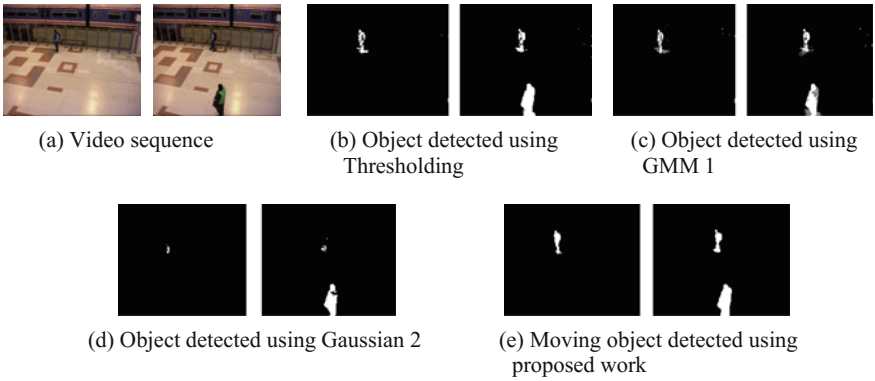
For object detection by means of manual thresholdng method [1], the object in an intended frame can be perceived by taking into consideration the absolute difference of background & intended frame and manual thresholding follows. Here, 0 to 255 is the range of threshold value. In the experiment different threshold values are considered.

First Gaussian mixture algorithm is based on two papers. Z. Zivkovic's "Improved adaptive Gaussian mixture model for background subtraction" in 2004 [14] and "Efficient Adaptive Density Estimation per Image Pixel for the Task of Background Subtraction" in 2006 [16]. A vital trait of the algorithm is that, for each pixel, it opts for the suitable number of Gaussian distribution. Second algorithm is a Gaussian Mixture-based Segmentation Algorithm. It was introduced in the paper "An improved adaptive background mixture model for real-time tracking with shadow detection" by P. KadewTraKuPong and R. Bowden in 2001 [15].

3.1 Qualitative Assessment

The foremost instance that is appraised in this analysis is PETS2006 video sequel. The video is taken from a railway station and depicts the area throughout a day. Two images of this video sequence are presented in Fig. 2a. The detected object by proposed work and three existing methods are given in Fig. 2b–d. From these outcomes, it is examined that manual thresholding [1], GMM-1 [14], and GMM-2 [15] based Background Subtraction techniques have recognized non-static objects from the frames with higher false rate. Compared to the three existing methods proposed method give better result.

The second example that is appraised in our analysis is a non-static background having one non-static object. Here we have chosen dynamic Water Surface video sequence. The video is attained in a less illumine situation with dynamic background because of variation in surface of water in the lake. Two actual video sequence of the video is presented in Fig. 3a. The object detected using thresholding, GMM-1, GMM-2are given in Fig. 3b–d, respectively. In Thresholding and



(a) Video sequence

(b) Object detected using Thresholding

(c) Object detected using GMM 1

(d) Object detected using Gaussian 2

(e) Moving object detected using proposed work

Fig. 2 PETS2006 video sequence

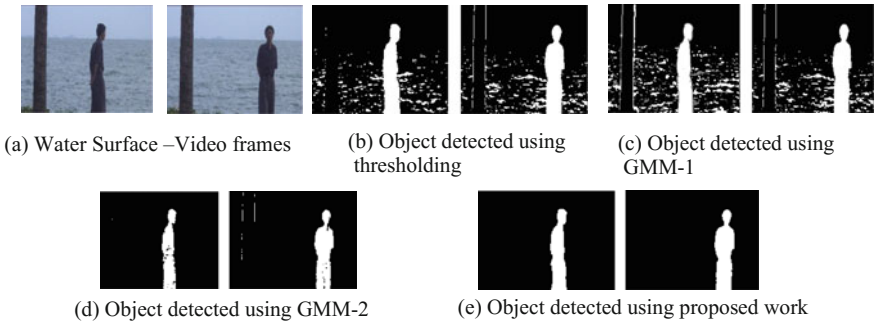


Fig. 3 Water surface

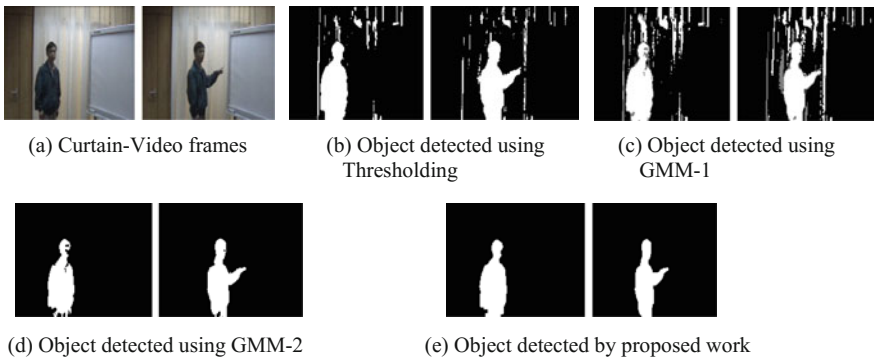
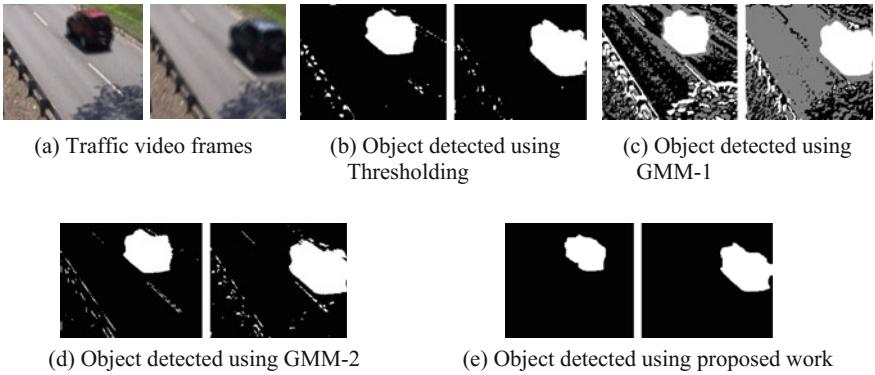


Fig. 4 Curtain video

GMM-1method moving water surface falsely detected as moving object. GMM-2 detects the moving object but the result is not accurate compared to proposed work.

The third example that is appraised in our analysis is Curtain video sequel. Non-static background and illumination variation are the features of this video sequence. Two original frames of Curtain video sequence is given in Fig. 4a. The non-static object detected using thresholding [1], GMM-1 [14], GMM-2 [15] are depicted in Fig. 4b–d appropriately. Output of proposed work is shown in Fig. 4e. Compared to existing three methods proposed work give accurate result. In Thresholding the moving background also detected as moving object. In GMM-1 some parts of moving object marked as shadow and moving curtain also detected as moving object. Compared to Thresholding and GMM-1, GMM-2 gives better result. But some part of object falsely detected as background. In our proposed scheme it correctly detected the moving object.



(a) Traffic video frames (b) Object detected using Thresholding (c) Object detected using GMM-1

(d) Object detected using GMM-2 (e) Object detected using proposed work

Fig. 5 Traffic video

Finally we are considered traffic video sequence. This video sequence is noisy and camera jitter also considered taking place in this video sequence. In our proposed work we are using spatio-temporal method so the problem of camera jitter can be solved. Because of camera jitter existing algorithms such as thresholding, GMM-1, and GMM-2 falsely detect the moving object. In the case of GMM-1 because of noise and poor picture quality most of the portion of background is detected as shadow. Proposed work give accurate result, so proposed work can solve camera jitter, poor image clarity problems (Fig. 5).

The proposed work is found to offer superior results in comparison with Thresholding, and two GMM-based approaches.

4 Conclusion and Future Work

In the paper, the proposal of a new spatio-temporal background subtraction scheme is projected. The proposed algorithm considers the spatial and temporal information to identify precise shape of non-static object from a video sequel. Local binary pattern gives the spatial information and Gaussian averaging gives the temporal information. In the projected method we initially construct a background model by using LBP histogram similarity. The moving object in the intended frame is recognized using Gaussian parameters by comparing it with background model constructed. From experimental outcomes, the projected proposal gives accurate result. The scheme provides accurate result for noisy scene, illumination variation, camera jitter and dynamic background. It is observed to outperform the avant-garde background subtraction schemes.

For future works we will consider shadow effect, camouflage problems. In proposed work we have given the constant manually, and taken an effort to find it automatically. Each pixel operations are independent and hence parallel implementation of this work will decrease the running time of the algorithm.

References

1. Bovic, A.L.: Image and Video Processing. Academic Press, New York (2000)
2. Lin, W., Sun, M.T., Poovendran, R., Zhang, Z.: Activity recognition using a combination of category components and local models for video surveillance. IEEE Trans. Circuits Syst. Video Technol. **18**(8), 1128–1139 (2008)
3. Piccardi, M., Jan, T.: Mean-shift background image modelling. In: Proceedings of International Conference on Image Processing, vol. 5, pp. 3399–3402 (2004)
4. Shi, Q., Cheng, L., Wang, L., Smola, A.: Discriminative human segmentation and recognition using semi-Markov model. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8 (2008)
5. Subudhi, B.N., Ghosh, S., Ghosh, A.: Moving object detection using Gaussian background model and Wronskian framework. In: Proceedings of 2nd International Conference on Advances in Computing, Communications and Informatics, pp. 1775–1780 (2013)
6. Heikkila, J., Silven, O.: A real-time system for monitoring of cyclists and pedestrians. In: Second IEEE Workshop on Visual Surveillance Fort Collins, Colorado, June 1999, pp. 74–81
7. Wren, C., Azarbayejani, A., Darrell, T., Pentland, A.: Pfinder: realtime tracking of the human body. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 780–785 (1997)
8. Stauffer, C., Grimson, W.: Learning patterns of activity using real time tracking. IEEE Trans. Pattern Anal. Mach. Intell. **22**(8), 747–767 (2000)
9. Stauffer, C., Grimson, W.E.L.: Adaptive background mixture models for real-time tracking. In: Proceedings of International Conference on Computer Vision and Pattern recognition, pp. 2246–2252 (1999)
10. Chan, A., Mahadevan, V., Vasconcelos, N.: Generalized Stauffer and Grimson background subtraction for dynamic scenes. Mach. Vis. Appl. **22**, 751–766 (2011)
11. Heikkila, M., Pietikainen, M.: A texture-based method for modeling the background and detecting moving objects. IEEE Trans. Pattern Anal. Mach. Intell. **28**(4), 657–662 (2006)
12. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell. **24**(7), 971–987 (2002)
13. Ahonen, T., Matas, J., He, C., Pietikäinen, M.: Rotation invariant image description with local binary pattern histogram fourier features. In: Salberg, A.-B., Hardeberg, J.Y., Jenssen, R. (eds.) SCIA 2009. LNCS, vol. 5575, pp. 61–70. Springer, Heidelberg (2009)
14. Zivkovic, Z.: Improved Adaptive Gaussian Mixture Model for Background Subtraction (2004)
15. KadewTraKuPong, P., Bowden, R.: An improved adaptive background mixture model for real-time tracking with shadow detection (2001)
16. Zivkovic, Z.: Efficient adaptive density estimation per image pixel for the task of background subtraction. Pattern Recogn. Lett. **27**(7), 773–780 (2006)

Design of CMOS Integrator Circuit for Sigma Delta ADC for Aerospace Application

Deepak Prasad and Vijay Nath

Abstract In this research paper, a design of CMOS integrator circuit for sigma delta ADC has been carried out. The proposed integrator claims a key role in low-power CMOS sigma delta ADC. The low-power CMOS sigma delta ADC is used to design smart temperature sensor for aerospace application which would sense a temperature range of −50 to 150 °C. The analog output of temperature sensor is digitized using low power CMOS sigma delta ADC. The integrator behaves as a low-pass filter for input signal and high-pass filter for quantization noise. Here, the input and output are measured across the capacitor. This circuit is designed using Cadence Virtuoso UMC90 nm CMOS technology. For the proper operation of the circuit, power supply is used in the range of +1.3 to −1.3 V. As the input square wave is applied, during the positive half cycle, the voltage across capacitor increases from zero to the maximum (peak value of applied voltage). During the negative half cycle, the capacitor starts to discharge and comes to zero.

Keywords Integrator · Sigma delta ADC · Capacitor · Resistor

1 Introduction

Sigma delta ADC is a low-cost, low-bandwidth, low-power, high-resolution ADC and it has varied applications in data acquisition, communications, sigma processing, and instrumentation. A sigma delta ADC comprises of a modulator part, which is analog in implementation and a decimal decimator stage. The modulator samples the input signal at much higher frequencies set by the oversampling ratio and converts the analog input signal into a pulse density modulated digital

D. Prasad (✉) · V. Nath
VLSI Design Group, Department of ECE, Birla Institute of Technology, Mesra, Ranchi, India
e-mail: prasaddeepak007@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_36

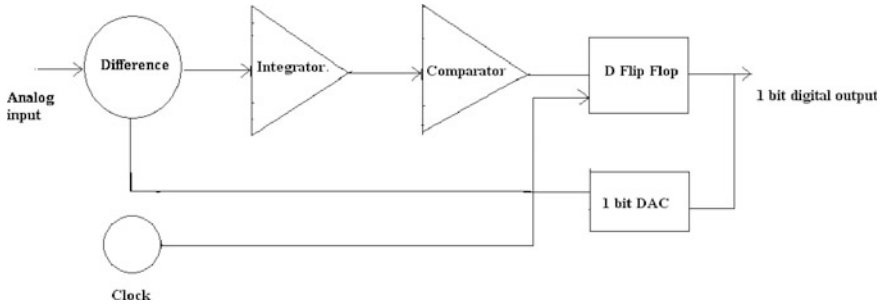
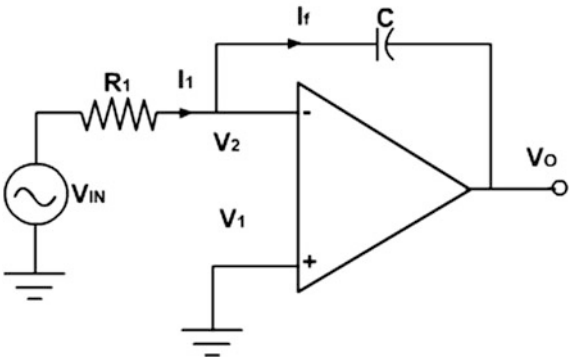


Fig. 1 Block diagram of sigma delta ADC comparator

Fig. 2 An integrator



containing both the original input signal and the unwanted out of band noise. The block diagram of sigma delta ADC comparator is shown in Fig. 1. The non-inverting and inverting continuous-time integrators have been designed using op-amps [1]. As its name implies, the CMOS integrator is an operational amplifier circuit that performs the mathematical operation of integration that is we can cause the output to respond to changes in the input voltage over time as the op-amp integrator produces an output voltage which is proportional to the integral of the input voltage [2]. Integrators, summer/subtractor (summing or difference amplifier) and scaling amplifiers are the important building blocks used in the design of continuous-time filters [3]. Figure 2 shows the basic circuit of an integrator.

The integrator will become unstable at very low frequency [4]. To eliminate the said drawback of instability, a resistance may be connected in parallel to capacitor (C). If square wave is given to input of the integrator then output comes in the form of Triangular wave. While spike waves input to an integrator gives output in the form of square wave. On the other hand, cosine wave output will be obtained when an integrator gets its input in terms of sine wave [5].

2 Methodology

A two-stage operational amplifier has been adopted in this paper to design an integrator. A capacitor has also been used who's charging and discharging property is used to give an output. The circuit provides an output voltage that is proportional to the time integral of the input. Here, CR is the integrator time constant [6]. Cadence implementation of op-amp-based integrator uses a 10 KΩ resistor and 10 pF capacitor in the negative feedback path. We apply a square wave input to port 'a' and get the integrated output at port 'Y' as shown in Fig. 3.

Voltage at the capacitor 'c' is

$$V_C = \mathrm{Q/c}, \tag{1}$$

where 'Q' is the charge at the capacitor

$$V_C = V_X - V_{OUT}$$
$$V_C = 0 - V_{OUT}$$

Differentiate both the sides with respect to 't'. We got

$$\frac{dV_C}{dt} = -\frac{dV_{OUT}}{dt} \tag{2}$$

Put the value of V_C in Eq. (2)

$$\frac{dQ}{cdt} = -\frac{dV_{OUT}}{dt}$$

$$-\frac{dV_{OUT}}{dt} = \frac{1}{c}\frac{dQ}{dt}$$

$$-\frac{dV_{OUT}}{dt} = \frac{1}{c}I_{IN}$$

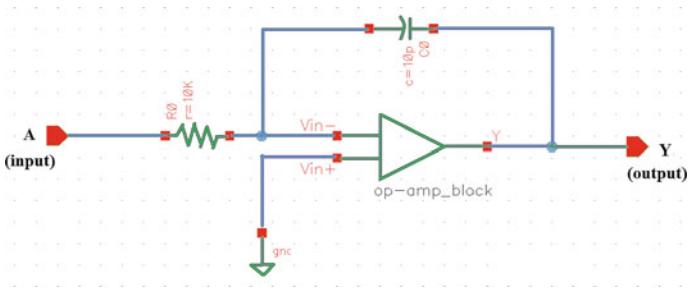


Fig. 3 Cadence implementation of integrated circuit

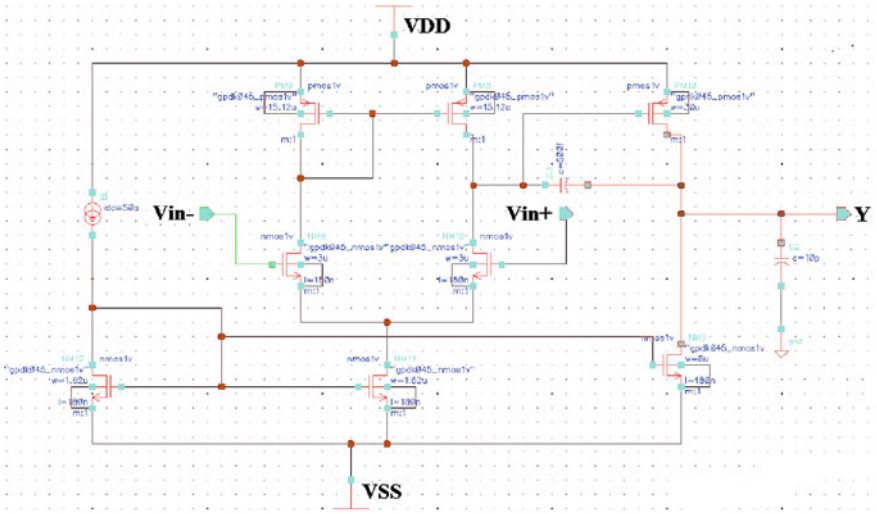


Fig. 4 Cadence schematic of op-amp circuit

Put the value of I_{IN}

$$-\frac{dV_{OUT}}{dt} = \frac{1}{c}\frac{V_{IN}}{r}$$

$$V_{OUT} = -\frac{1}{rc}\int_0^t V_{IN}dt$$

The main ingredient of ADC is high-gain operational amplifier which has been designed in Fig. 4. To let the signal swing above and below the ground, a power supply of VDD and VSS is used.

A biasing voltage of 1 V, supply voltage VDD of +1.3 V and VSS of −1.3 V and 1 kHz frequency has been assumed to simulate the design. A 48.45 dB open loop gain has been observed for the designed operational amplifier. The open loop gain is shown in Fig. 5 and frequency response is shown in Fig. 6.

3 Simulation and Result

When a square wave is given as input to an integrator amplifier, then the capacitor will charge and discharge in response to changes in the input signal. This results in the output signal being that of a sawtooth waveform whose frequency is dependent upon the RC time constant of the resistor/capacitor combination. In this circuit, we apply +500 to −500 μV peak-to-peak voltage to the input (red color) of the

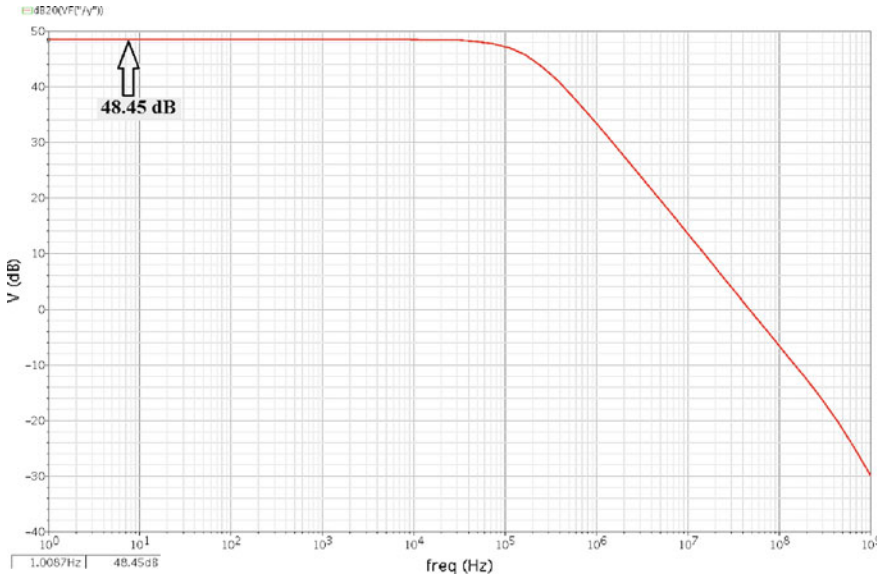


Fig. 5 Cadence output for op-amp circuit

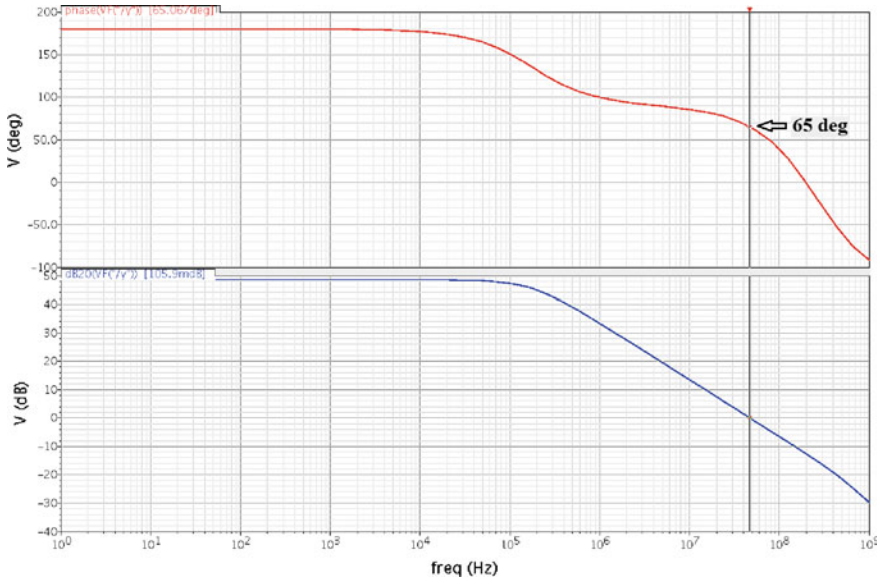


Fig. 6 Frequency response of op-amp

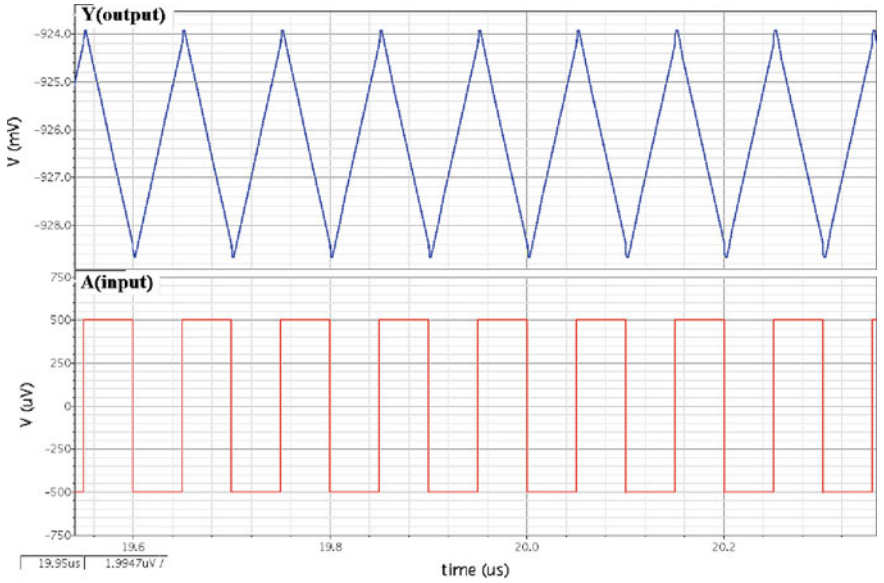


Fig. 7 Cadence output of integrator circuit

integrator. For positive input, capacitor starts charging and for negative input, capacitor starts discharging. Hence, we get a triangular waveform at the output (blue color). See Fig. 7.

4 Conclusion

It has been observed that the circuit which is designed has worked properly. For op-amp, its output stability is very good for 48.45 dB gain. For integrator, we got charging and discharging pulse at the output. For the given rectangular pulse, we got the triangular pulse.

Acknowledgements We are thankful to RESPOND ISRO, Bangalore for funding this research work. We are also thankful to Prof. V. R. Gupta, HOD ECE, Prof. T. Ghosh, Dean (SR), and Prof. R. Sukesh Kumar, Dean (FA) & Prof. M.K. Mishra, Vice chancellor, BIT Mesra Ranchi for providing infrastructure facility to carry out this research work.

References

1. Allen, P.E., Holberg, D.R.: CMOS Analog Circuit Design. Oxford University Press (2002)
2. Weste, N.H., Eshraghian, K.: Principles of CMOS VLSI Design, vol. 2. Addison Wesley, MA (1993)

3. Kamath, D.V.: Overview of OPAMP and OTA based integrators. Int. J. Innov. Res. Electr. Electron. Instrum Control Eng. **3**(9), (2015)
4. Sedra, A., Smith, K.C.: Microelectronic Circuits. Oxford University Press (2004)
5. Shem, B., Kozak, M., Friedman, G.: A high-speed CMOS op-amp design technique using negative miller capacitance. In: IEEE (2004)
6. Mandal, P., Visvanathan, V.: A new approach for CMOS op-amp synthesis. In: VLSI Design, Twelfth International Conference (1999)

Aspects of Low-Power High-Speed CMOS VLSI Design: A Review

Prolay Ghosh, Tanusree Saha and Barsha Kumari

Abstract VLSI and nanocomputing has become the most desirable feature of any integrated chip. Computers are also becoming more portable. ICs are being introduced everywhere. This huge implementation of IC has also opened a scope of research. Size of IC has become an issue to think about. Power dissipation is also another important consideration as performance of VLSI chip design. Low-power high-speed CMOS circuit design methodologies will be elaborated in this paper.

Keywords Power dissipation · Threshold voltage · Leakage current · Propagation delay · VTCMOS · MTCMOS

1 Introduction

Digital systems are implemented using very large number of logic gates. For space and other economic consideration, it is desirable to implement the system with as few integrated circuits chips as possible. It follows that one must pack as many logic gates as possible on an IC chip. At present, around 1,00,000 or more gates can be fabricated on a single IC chip in what is known as VLSI. To keep the power dissipation in the chip to acceptable limits, the power dissipation per gate must be kept to a minimum. Indeed, a very important measure of the logic circuits is the power it dissipates. During the introduction of VLSI technology, it was completely focused on execution time of real-time function on an IC. So the designer's concentration was on various aspects of speed optimization. As a result, designers have integrated semi-

P. Ghosh (✉) · T. Saha · B. Kumari
Department of Information Technology, JIS College of Engineering, Kalyani, India
e-mail: prolayghosh.jis@gmail.com

T. Saha
e-mail: tanusreesah@gmail.com

B. Kumari
e-mail: barsha.kumari15@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_37

conductor ICs with various complex signal processing and graphical processing units but we sacrificed the portability. In those IC, we were able to execute a complex problem with a high execution rate but portability was not improved. But nowadays tablets, smart phones, etc., are capturing the electronics market very fast, where not only speed but portability also has become a important factor. Only way to increase the portability is to reduce the chip size. On the other hand, number of transistors on a chip is getting increased day by day. As VLSI technology is improving, demand of system on chip is also increasing. As a result of all these, the chip is leading to high power dissipation. Consumer demands a portable electronic device, avails all the features, generation of heat is less, fast, and consumes less power. Here, we have discussed various components about the power dissipation and represent some existing methodologies to reduce the power dissipation.

2 Moore's Law

In the year of 1960s, Gordon Moore presented the Moore's law [1], where the number of transistors counts on a chip would become double every 18 months. Following the rule over the past 30 years, the number of transistors per chip has doubled about once a year. The following table gives a real example of Moors Law.

Microprocessor	Year	No of transistors	Feature size (μ)
80286	1982	1,34,000	1.5
80386	1985	2,75,000	1.5
80486	1989	12,00,000	1.0
Pentium	1993	31,00,000	0.8
Pentium pro	1995	55,00,000	0.6

3 MOSFET

MOSFET is a three-terminal device. The basic principle involved is that voltage is applied to one terminal (Gate) and that gate voltage controls the flow of current between two terminals (Source to Drain terminal). Threshold voltage is that amount of voltage initiates current will flow from source to drain. That means current will flow only when supply voltage crosses the threshold voltage. The threshold voltage for pMOS is negative and for nMOS is positive.

In MOSFET, when the input changes the output does not change instantly. For instance in a CMOS inverter at T_1 time input changes from low to high but the output does not change at T_1 but it takes very small time to change from high to low and changes at T_2. This T_2-T_1 delay is called propagation delay. This delay affects the performance of the circuit. As output changes from low to high or high to low,

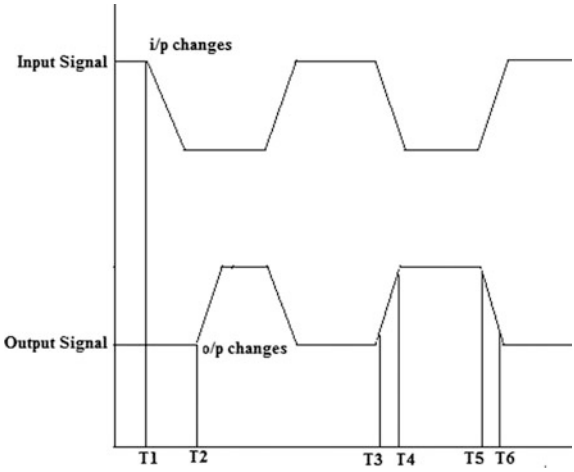


Fig. 1 T2-T1 => Propagation delay, T4-T3 => Rise time, T6-T5 => Fall Time

initiation of change and completion of change do not happen at the same point. Difference between this two time points is called rise time or fall time.time (Fig. 1).

$$\text{Time of slope} = (\text{Rise time} + \text{Fall time})/2$$

4 Power Dissipation

Power dissipation has three components.

4.1 Static Power

The semiconductor consumes static power dissipation due to leakage current. Leakage current is that current which flows when the ideal current is zero. When a voltage is applied to a capacitor, current is conducted through the capacitor that means the capacitor is on. The current induction causes some power dissipation. As soon as the applied voltage is withdrawn, the capacitor should reach to the off stage. But the capacitor characteristic is that it discharges the voltage very slowly after eliminating the applied voltage. For that period of time, small amount of current is conducted. It caused some unintended power loss from a capacitor. Leakage power is dissipated when the device is actually is in inactive stage. It is observed that in deep submicron technology, leakage current contributes 50% of total power consumption.

$$\text{Static power dissipation } (P_{static}) = \text{Supply Voltage (VDD)} \times \text{Leakage Current } (I_{leakage}) \quad (1)$$

Two major components of Leakage current are:

(i) Subthreshold leakage: Current conduction takes place when supply voltage is grater then the threshold voltage of the MOSFET. But practically, a small amount of current is conducted between source and drain in MOS transistor through gate voltage below the threshold voltage current that caused sub-threshold leakage. It can be as high as 80–90% of total circuit leakage.
(ii) Gate oxide leakage: If a high electric field is applied to a MOS transistor with low oxide thickness then a current flows through the oxide tunnel called gate oxide leakage.

4.2 Dynamic Capacitive Power

CMOS digital circuits consume power while charging and discharging various load capacitances. Current conduction does not start as soon as the supply voltage is applied. The output node capacitance takes some time to get charged [2]. During this charge up phase, the energy drawn from the power supply is dissipated as heat in the conducting pMOS transistor. Again when supply voltage drops to 0 means no energy is drawn from power supply, energy stored in the output capacitance during charge up is dissipated as heat in the conducting nMOS transistors as the capacitor discharges slowly.

$$P_{dynamic} = \text{Load capacitance (CL)} \times \text{VDD2} \times \text{operating frequency (Fop)} \quad (2)$$

The dynamic capacitive power dissipation is based on the assumption that the output node undergoes one power-consuming transition (0 to VDD) in each clock cycle. This assumption is not always correct. Depending on the circuit topology, logic style, and the input signal statistics, the node transition rate can be slower than the clock rate. So, we introduce node transition factor or activity factor (α), which is effective number of power-consuming voltage transitions experienced per clock cycle. It is directly proportional to dynamic power consumption. So the previously stated equation of dynamic capacitive power is modified as (Fig. 2)

$$P_{dynamic} = \alpha \times CL \times VDD2 \times Fop \quad (3)$$

4.3 Dynamic Short-Circuit Power

Short-circuit power is caused by the short-circuit currents that arise when pairs of pMOS/nMOS transistors are conducting simultaneously. As the input voltage makes a transition from low to high, ideally pMOS should be turned off and nMOS

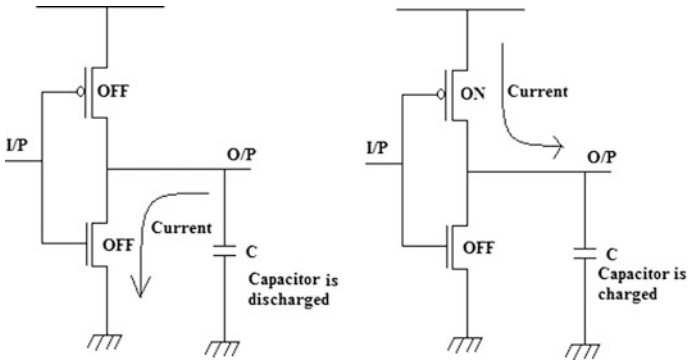


Fig. 2 Current flow through MOSFET at off stage

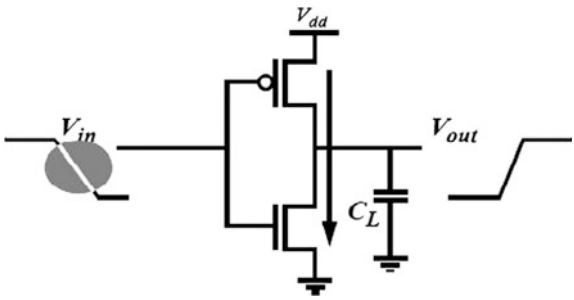


Fig. 3 Dynamic short-circuit power

should be turned on. But the process takes small time till then both the transistors remain on and it provides a short-circuit path from supply voltage to ground. This small time is called time of slope (Rise time/Fall time) (Fig. 3).

$$\text{Short} - \text{circuit power (Psc)} = \text{VDD} \times \text{time of slope (Tsc)} \times \text{Imax} \quad (4)$$

5 Power Dissipation Reduction Techniques

There are various power dissipation reduction techniques that are discussed below

5.1 Lowering the Capacitance

By downsizing driver and receiver gate power consumption can be reduced. The length of the wire can be minimized to reduce load capacitance by placing the

transistors properly and routing them properly. Decreasing capacitance is also difficult as it needs to scale down the device and wiring.

5.2 Lowering the Operating Frequency

Higher frequency encourages better performance but it causes high power consumption. To reduce frequency, parallel processing can be introduced but it increases hardware overhead and requires extrusive change at an architecture or algorithm design.

5.3 Lowering the Supply Voltage

From the Eqs. (1), (3) and (4), it is clear that supply voltage and power dissipation have quadratic dependency. So, the most acceptable choice is to reduce the supply voltage to reduce the power dissipation. But it is proven that if supply voltage is reduced then circuit delay will increase. We have designed CMOS NAND using ORCAD P-Spice [3] (Fig. 4). The NAND gate simulated with the supply voltage 5 V and then with the supply voltage 2 V. In Fig. 5, supply voltage is 5 V delay is near about 0.9 ms. In Fig. 6, supply voltage is 2 V delay is nearly 2.8 ms. So 63.33% increment in delay is observed. It also affects the speed of the circuit.

5.3.1 Variable Threshold Voltage (VTCMOS)

VTCMOS technique reduces the supply voltage by varying the threshold voltage. When supply voltage is applied, current will flow through the critical path and through noncritical path no current is conducted. So for critical paths, threshold voltage is lowered in such a way that the MOSFET switching speed increases [4]. But for noncritical path, a high threshold voltage is applied to reduce the leakage current. In standby mode, the control circuit generates a lower substrate voltage for NMOS and higher substrate voltage for PMOS. This causes the threshold voltage of both the NMOS and PMOS to increase, thus reducing the leakage current (Figs. 5 and 6).

5.3.2 Multiple Threshold Voltage (MTCMOS)

Low threshold voltage and reduction of leakage current both are important for low-power high-speed circuit. As we have seen that by lowering the threshold voltage, we can reduce the power dissipation but it increases the standby current. So N-channel and P-channel MOSFET with two different threshold voltages is employed in a single

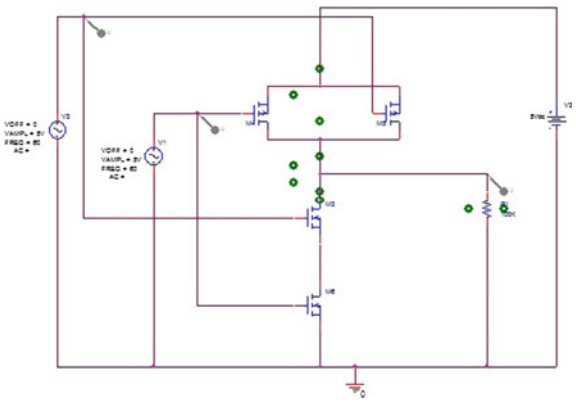


Fig. 4 CMOS NAND gate implemented using ORCAD P-Spice

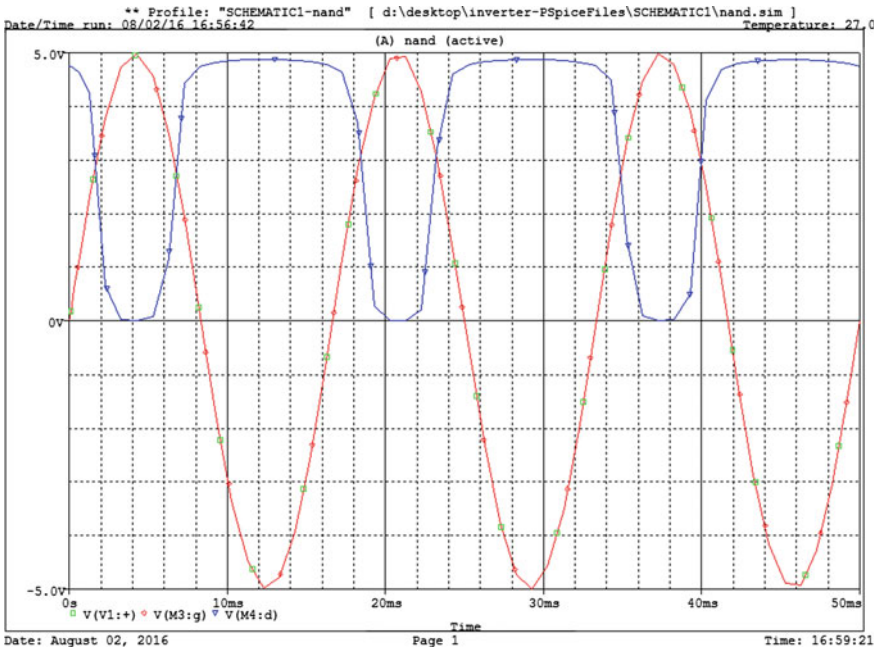


Fig. 5 CMOS NAND gate output timing diagram for supply voltage 5 V

chip with two different operational modes 'active' and 'sleep'. Low threshold voltages are placed in the critical path while MOS transistors with high threshold voltage are placed in the noncritical path. In the critical path, gate delay time should be as low as possible [4]. Hence, delay reduction is traded off against leakage by placing low threshold transistors causing faster gate transitions. However, in the non-timing critical

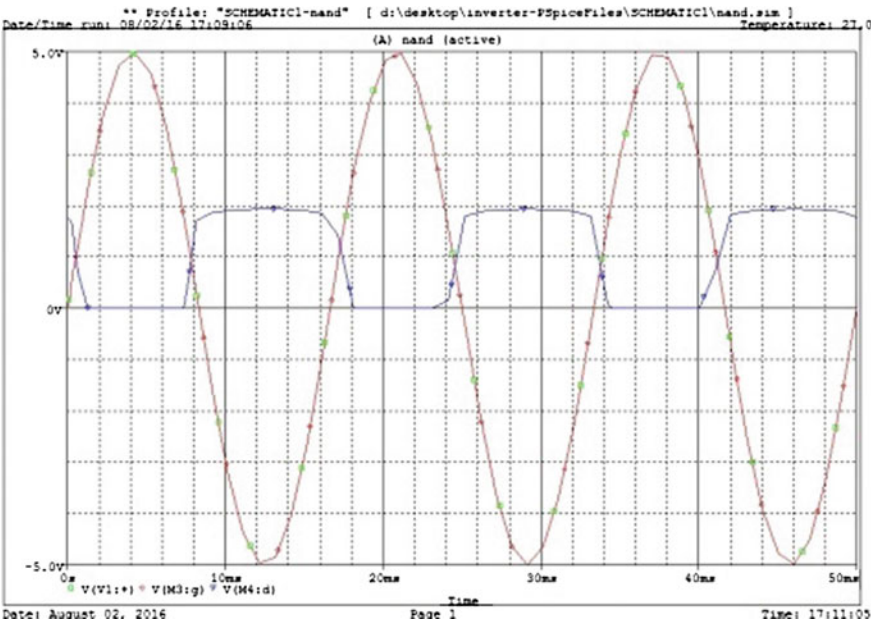


Fig. 6 CMOS NAND gate output timing diagram for supply voltage 2 V

path, high threshold transistors are placed to reduce leakage as speed/delay has less importance.

5.3.3 Dynamic Threshold Voltage (DTCMOS)

In dynamic threshold CMOS (DTMOS), the threshold voltage is altered dynamically to suit the operating state of the circuit. To reduce leakage current in the standby mode high threshold voltage is applied, while in the active mode of operation, a low threshold voltage is applied to ensure high current conduction.

5.4 Lowering the Activity Factor

The toggling rate of a signal with respect to clock is termed as activity factor. Dynamic power dissipation is directly proportional to activity factor.

5.4.1 Clock Gating

The most popular technique to reduce activity factor is clock gating. The main logic behind this technique is to disconnect the clock from the device when the data going into the device is not going to change. When EN = 1, clock can be applied to

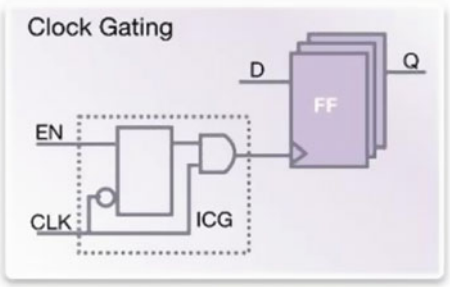


Fig. 7 Clock gating

functional block through AND gate. However, to avoid unnecessary switching when the functional block is inactive, the EN input can be made 0 so as to disable the clock propagation when not required [2]. This causes the clock to become totally inactive when the functional block is in nonactive mode, thus eliminating any sort of activity in the circuit (Fig. 7).

5.5 Lowering the Rise Time, Fall Time, and Propagation Delay

Rise and fall time can be reduced by reducing the gate length of the MOSFET. Strained silicon MOSFET has also less rise and fall time than bulk silicon MOSFET. As electrons can move 70% faster allowing strained silicon transistors to switch 5% faster.

The short-circuit power can also be reduced by

i. Device size reduction and
ii. Increasing threshold voltage

Both of which is also used for leakage reduction.

6 Conclusion

The need of lower power system is being driven by many market segments. Unfortunately, designing for low power adds another dimension to the already complex design problem and the design has to be optimized for power as well as performance and area. In this paper, we have tried to discuss various aspects of power dissipation and various methods to reduce the power dissipation as well as increase the performance. But as the need of market is changing regularly, the need of designing smaller chip is also increasing. That need leads to the era of VVLSI and ULSI. Well-planned physical design and floor planning are very important for the reduction of chip size. At the same time, designer needs to consider the power dissipation and performance of the chip.

References

1. http://download.intel.com/newsroom/kits/ml50/pdfs/GordonMoore_QA.pdf
2. Sivakumar, R., Jothi, D.: Recent trends in low power VLSI design. IJCEE **6**, 869 (2014)
3. http://www.physics.ntua.gr/~yorgos/ilektronika/PSPICE.pdf
4. Kuroda, T.: Low power, high speed CMOS VLSI design. In: Proceedings of the 2002 IEEE International Conference on Computer Design: VLSI in Computers and Processors (ICCD'02)
5. Fadl, O.S., Abu-Elyazeed, M.F., Abdelhalim, M.B., Amer, H.H., Madian, A.H.: Accurate dynamic power estimation for CMOS combinational logic circuits with real gate delay model. J. Adv. Res. **7**, 89–94 (2016)
6. Reuben, J., Mohammed, Z.V., Kittur, H.M., Shoaib, M.: A buffer placement algorithm to overcome short-circuit power dissipation in mesh based clock distribution network. Eng. Sci. Technol. Int. J. **18**, 135–140 (2015)
7. Koshya, L.M., Chandran, G.J.: Low leakage and high performance tag comparator implemented in 180 nm CMOS technology. Proc. Comput. Sci. **46**, 1261–1267 (2015)
8. Sanadhyaa*, M., Kumar, M.V.: Recent development in efficient adiabatic logic circuits and power analysis with CMOS logic. Proc. Comput. Sci. **57**, 1299–1307 (2015)
9. Lina, G.J.Y., Kuo, J.B.: Novel power consumption reduction strategy (PCRS) using mixed-VTH cells for optimizing the cells on critical paths for low-power SOC. IERI Proc. **4**, 231–236 (2013)
10. De, B.P., Kar, R., Mandal, D., Ghoshal, S.P.: Optimal CMOS inverter design using differential evolution algorithm. J. Electr. Syst. Inf. Technol. **2**, 219–241 (2015)
11. Upadhyay, P., Kar, R., Mandal, D., Ghoshal, S.P.: A low power CMOS voltage mode SRAM cell for high speed VLSI design. IEEE Xplore (2013)
12. Kulkarni, N., Yang, J., Seo, J.-S., Vrudhula, S.: Reducing power, leakage, and area of standard-cell ASICs using threshold logic flip-flops. IEEE Trans. Very Large Scale Integr. (VLSI) Syst. 2873–2886 (2016)
13. Mirmotahari, O., Berg, Y.: Reliability of high speed ultra low voltage differential CMOS logic. circuits and systems. Sci. Res. **6**, 121–135. doi:10.4236/cs.2015.65013
14. Kumar, P., Sharma, R.K.: Low voltage high performance hybrid full adder. Eng. Sci. Technol. Int. J. **19**(1), 559–565 (2016)

An Intelligent Beta Reputation and Dynamic Trust Model for Secure Communication in Wireless Networks

S. Sathish, A. Ayyasamy and M. Archana

Abstract Networking technology is playing a major role in daily life for communicating each other. Security is the most important issue in wireless networks. Recently, trust and reputation mechanisms are used for providing security through monitoring the behaviour. However, the existing works lack in providing reliable security to wireless networks. In this paper, we propose an intelligent beta reputation and dynamic trust model (IBRDT) for providing security in wireless networks. This model is the combination of dynamic trust and beta reputation trust for secure routing in wireless networks.

Keywords Wireless networks · IBRDT · Secure routing · Beta reputation · Dynamic trust model

1 Introduction

Wireless networks are useful for sharing knowledge and information mutually in this real world. Rapid growth of this computer network utilization and security issues are also increasing simultaneously due to the huge number of users. Moreover, wireless networks consist of a large number of nodes without physical connection. The applications of wireless networks such as medical, computer security, defense and surveillance are useful and essential component of human life. Moreover, the data gathered by the nodes are forwarded and routed to the base station either directly or through neighbouring nodes. In such a scenario, each node

S. Sathish (✉) · A. Ayyasamy · M. Archana
Faculty of Engineering and Technology, Department of Computer Science and Engineering, Annamalai University, Annamalai Nagar, Tamil Nadu, India
e-mail: sathish.srajan@gmail.com

A. Ayyasamy
e-mail: samy7771@yahoo.co.in

M. Archana
e-mail: archana.aucse@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_38

of the network is capable of providing service within their transmission range. In wireless networks, data are gathered from nodes and are sent to the base station which is called as the sink node.

In wireless network, trust specifies the reliability or trust worthiness of node. Trust mechanism can be classified in different ways based on how the trust values are calculated. Trust is subjective based on individual node behaviour in a group or network. Trust mechanism is classified into two: namely direct and indirect trust mechanisms [1]. Direct trust is considered as a basic opinion about the particular node. Indirect trust is considered as a second opinion which is collected from some other nodes that are located as neighbour. Direct trust values are calculated between nodes. Indirect trust values are calculated between the node and neighbour nodes. Dynamic trust mechanism is helpful to know the current trust value of the particular node in ad hoc networks [2]. Moreover, trust and reputation are multidisciplinary concepts with different definitions and evaluations in various fields [3, 4].

In this paper, a new intelligent beta reputation and dynamic trust model (IBRDT) is proposed and implemented for effective communication in wireless networks. The proposed system calculates the dynamic trust value by using the direct trust and beta reputation trust for providing secure communication.

The rest of the paper is organized as follows: Sect. 2 discusses various secured routing protocols proposed by many authors in the past. Section 3 explains the proposed work. Section 4 provides the results and discussion. Finally, Sect. 5 gives conclusion and future works.

2 Literature Survey

Josang and Ismail [4] made a survey of trust and reputation management systems in wireless communications. They discussed about the latest methods which have been proposed by researchers to manage the trust and reputation in wireless communication systems. Zhu et al. [1] proposed an authenticated trust and reputation calculation and management system for cloud and sensor networks. Anupam Das and Mohammad Mahfuzul Islam [2] proposed a dynamic trust computation model for secured communication using agents. Fenye Bao et al. [5] proposed a hierarchical trust management scheme which is the most useful for detecting intruders through effective secured routing in wireless sensor networks.

Geetha and Jayakumar [6] proposed a new free roaming mobile agent model which is based on reputation mechanism for secure communication. They generated a trust value based routing table and double encryption mechanism for effective secure communication in networks. They used mobile agent for ensuring the confidentiality, integrity, authenticity and reliability. Younghun Chae et al. [7] discussed about the vulnerabilities of existing redemption schemes, and also described a new trust management and redemption scheme that discriminates between temporary errors and disguised malicious behaviours with a flexible design. They also analysed the experimental results trust management scheme and

also demonstrated the advantages of the proposed scheme with simulation conducted in a wireless sensor network (WSN). Seyed Ali Mousavifar and Cyril Leung [8] proposed a collaborative spectrum sensing protocol for improving the energy efficiency of the traditional collaborative spectrum sensing protocol.

Zhu et al. [1] explored the authentication as well as trust and reputation calculation and management of CSPs and SNPs, which are two very critical and barely explored issues with respect to CC and WSNs integration. Moreover, they proposed a novel trust-based authentication system for wireless networks. Stylianos Kraounakis et al. [9] proposed a computational model for trust establishment based on a reputation mechanism for effective communication. Muneeswari et al. [10] proposed an intelligent data gathering and energy efficient routing algorithm for mobile wireless sensor networks. Logambigai and Kannan [11] proposed a fuzzy logic based unequal clustering for WSNs.

Al-Jarrah et al. [12] have proposed Randomized Data Partitioned Learning Model (RDPLM) relying on a compact network feature set and feature selection techniques, simplified subspacing and a multiple randomized metalearning technique. Rajeswari et al. [13] have proposed efficient mechanism for the detection of malicious nodes in MANET that uses back-off duration to avoid the formation of too many clusters unnecessarily at the initial stage, which reduces the energy consumption.

3 Proposed Work

In this work, an intelligent trust model called intelligent beta reputation and dynamic trust (IBRDT) is proposed for effective secure communication. A widely used way to map the observed information from the evidence space to the trust space is the beta distribution. Let *s* and *f* represent the total amount of positive and negative feedbacks in the evidence space about target entity, then the trust worthiness *t* of a subject node is then computed as,

$$\mathrm{t} = \mathrm{s} + 1/\mathrm{f} + \mathrm{s} \tag{1}$$

$$\mathrm{DyT} = \mathrm{Dynamic\ Trust}\,(\mathrm{t}, <\mathrm{t1}, \mathrm{t2}>) \tag{2}$$

IBRDT is the combination of Dynamic Trust (DyT) and Beta Reputation Trust. Intelligent beta reputation model is used for calculating the beta direct trust value using intelligent agents. Here, the intelligent agents are used for monitoring the node trust during particular time duration dynamically. The proposed intelligent beta reputation system demonstrates the behaviours of each individual node as a binary event. This binary event is modelled by the beta distribution which is commonly used to represent the posterior probability of a binary event using intelligent agents. Dynamic trust and reputation of each node is evaluated by the features provided by the beta distribution that acts as a basis. The beta family of probability density functions (PDFs) is a set of continuous function indexed by two

parameters α and β. In beta reputation system, α is assigned as the number Np of positive ratings plus 1 and β is assigned as the number Nn of negative ratings plus 1. Initially, dynamic trust is the expectation of positive behaviour from a node. In future interactions, the trust worthiness value is calculated as,

$$\frac{\alpha}{\alpha+\beta} \equiv \frac{N_P+1}{N_P+N_n} + \mathrm{DyT} \tag{3}$$

P represents the decay factor or forgetting can be applied to assign more weight to new ratings and gradually the older ratings are decreased. Intelligent beta reputation and dynamic trust value is calculated as follows,

$$IBRDT = \frac{S+1}{F+S+2} + \frac{dS+1}{dF+dS} + DYT \tag{4}$$

IBRDT is the combination of dynamic trust and beta reputation trust. The proposed intelligent beta reputation model is used for calculating the beta trust value dynamically. The proposed work consists of a trust-based secure routing algorithm that works in three phases namely trust score evaluation, threshold setting, and routing based on the trust values. This proposed work focuses on two important aspects namely beta reputation and dynamic trust based secure routing. The trust-based secure routing algorithm is the main focus of this work. The steps of the proposed secured routing algorithm are as follows:

Dynamic trust based secure routing algorithm

Step 1: Let T_v (n_1, n_2...n_m) = 0. // T_v indicate Trust value, n_1, n_2...n_m are nodes.
Step 2: Every node (n_1, n_2...n_m) are considered as source node in different time duration (t_1, t_2)
Step 3: Send messages to the neighbour nodes
Step 4: HC = HC + 1
Step 5: Start the Scheduler Class to execute the simulation.
Step 6: If it received the request from neighbour nodes then
ensure that the node is destination node
Else If it is destination then
It sends the acknowledgement to its neighbouring nodes.
Step 7: Compute the trust score for all the nodes using Eq. 1.
Step 8: Compute the dynamic trust score for all the nodes using Eq. 2.
Step 9: Compute the overall trust score for all the nodes using Eq. 3.
Step 10: If Minimum value (Tkc) < Threshold then
Detect the malicious node
Else
Update the routing table with new node.
Step 11: Perform routing performance.

The proposed secure routing algorithm calculates the trust value dynamically. The trust values are calculated during different time intervals for all the participant nodes of the network scenario. The participant nodes ensured the proper destination node by receiving the acknowledgement for their messages. Similarly, the trust score and dynamic trust score have been calculated for the individual nodes using the Eqs. (2) and (3). Threshold values are fixed by the intelligent agents and checked with the dynamic trust scores of all nodes in the network scenario. If the dynamic score of the particular node is less than the threshold value, then the particular node must be considered as malicious node and it is also avoided for performing routing. Finally, the routing process is performed with all other nodes which are having the dynamic scores above the threshold.

4 Results and Discussion

We have implemented the proposed routing algorithm using NS2 (Version 2.34.1) by using the existing AODV routing protocol. The topology of the wireless network depends on the pause time and mobility speed and also it changes its topology frequently when pause time is less and mobility speed is more. The performance of AODV protocol in presence of malicious node is compared with the performance of proposed technique in this work. Figure 1 describes the trust score variation between the existing and proposed system. From Fig. 1, it can be seen that the proposed system performs well than the existing system. This is due to the use of intelligent reputation mechanism and dynamic trust value calculation.

Figure 2 shows the delay analysis of the proposed system and the existing AODV protocol. From Fig. 2, it can be observed that the performance of the proposed system is better than the existing protocol in terms of delay.

Figure 3 shows the packet drop ratio analysis of the proposed routing algorithm and the existing AODV.

From Fig. 3, it can be observed that the packet drop ratio gradually decreases in this proposed IBRDT when it is compared with AODV with the minimum number

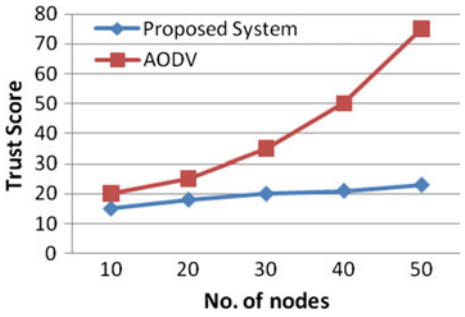


Fig. 1 Average trust score analysis in percentage

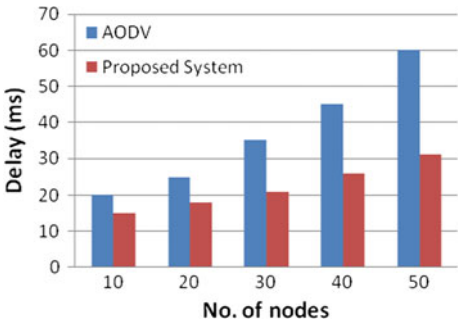


Fig. 2 Delay analysis

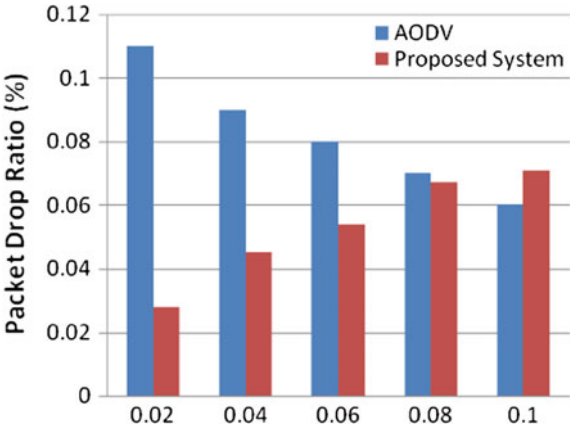


Fig. 3 Packet drop ratio analysis

of malicious nodes are present in the network. This is due to the use of intelligent agent, dynamic trust and the beta reputation system.

Figure 4 shows the comparative analysis between the proposed system and the existing systems.

From Fig. 4, it can be observed that the performance of the proposed system is better than the existing systems. This is due to the use of dynamic trust and reputation system in the proposed system.

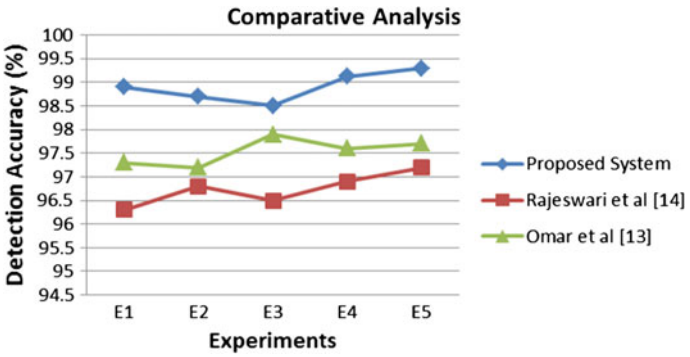


Fig. 4 Comparative analysis

5 Conclusion

An intelligent beta reputation and dynamic trust model is proposed and implemented for effective secure communication. Moreover, an intelligent secure routing algorithm has been proposed, discussed and implemented in this research work. From the experiments conducted using this secure routing algorithm, it has been shown that the trust and reputation calculation and management for secure communication in wireless networks.

References

1. Zhu, C., et al.: An authenticated trust and reputation calculation and management system for cloud and sensor networks integration. IEEE Trans. Inf. Forensics Secur. **10**(1), 118–131 (2015)
2. Das, A., Islam, M.M.: Secured trust: a dynamic trust computation model for secured communication in multi agent systems. IEEE Trans. Dependable Secure Comput. **9**(2), 261–274 (2012)
3. Ganeriwal, S., Balzano, L.K., Srivastava, M.B.: Reputation-based framework for high integrity sensor networks. ACM Trans. Sens. Netw. **4**(3) (2008)
4. Josang, A., Ismail, R.: The beta reputation system. In: Proceedings of 15th Bled Electronic Commerce Conference, 2002, pp. 324–337
5. Bao, F., Chen, I.-R., Chang, M., Cho, J.-H.: Hierarchical trust management for wireless sensor networks and its applications to trust-based routing and intrusion detection. IEEE Trans. Netw. Serv. Manage. **9**(2), 169–183 (2002)
6. Geetha, G., Jayakumar, C.: Implementation of trust and reputation management for free-roaming mobile agent security. IEEE Syst. J. **9**(2), 556–566 (2005)
7. Chae, Y., Dipippo, L.C., Sun, Y.L.: Trust management for defending on-off attacks. IEEE Trans. Parallel Distrib. Syst. **26**(4), 1178–1191 (2015)
8. Mousavifar, S.A., Leung, C.: Energy efficient collaborative spectrum sensing based on trust management in cognitive radio networks. IEEE Trans. Wireless Commun. **14**(4), 1927–1939 (2015)

9. Kraounakis, S., Demetropoulos, I.N., Michalas, A., Obaidat, S.M., Sarigiannidis, P.G., Louta, M.D.: A robust reputation-based computational model for trust establishment in pervasive systems. IEEE Syst. J. **9**(3), 878–891 (2015)
10. Muneeswari, S.J., Ganapathy, S., Kannan, A.: Intelligent data gathering and energy efficient routing algorithm for mobile wireless sensor networks. Asian J. Inf. Technol. **15**(5), 921–927 (2016)
11. Logambigai, R., Kannan, A.: Fuzzy logic based unequal clustering for wireless sensor networks. Wireless Netw. **22**(3), 945–957 (2016)
12. Al-Jarrah, O.Y., Alhussein, O., Yoo, P.D., Muhaidat, S., Taha, K., Kim, K.: Data randomization and cluster-based partitioning for Botnet intrusion detection. IEEE Trans. Cybern. **46**(8), 1796–1805 (2016)
13. Rajeshwari, A.R., Kulothungan, K., Ganapathy, S., Kannan, A.: Malicious nodes detection in MANET using back-off clustering approach. Circuits Syst. **7**, 2070–2079 (2016)

Effect of Incidence Angle on Optical Bandwidth in Ternary Photonic Crystal for Filter Application

Romi Dey, Meenakshi Banerjee, Antara Das and Arpan Deyasi

Abstract Optical bandwidth of ternary photonic crystal based Butterworth filter is computed for polarized incidence of electromagnetic wave; and effect of incidence angle and structural parameters are studied within lower range on the filter performance. Result is compared with that obtained for normal incidence. Transfer matrix technique is adopted for calculation; and SiO_2/air/TiO_2 material system is considered for simulation purpose. Simulated findings in favor of p-polarized wave incidence for varying incidence angle owing to higher bandwidth and less ripple in passband for filter application.

Keywords Ternary photonic crystal · Optical bandwidth · Angle of incidence · Butterworth filter · Structural parameters

1 Introduction

Photonic crystal may be considered as a Bragg grating structure where refractivity of the constituent materials becomes a periodic function of space coordinate along the direction of electromagnetic wave propagation [1, 2]. It exhibits photonic bandgap; which can restrict propagation of electromagnetic wave upto a particular region of wavelength, and allow the other spectra [3]. The extent of bandgap depends on the refractive indices of the constituent materials [4, 5], thicknesses of

R. Dey (✉) · M. Banerjee · A. Das · A. Deyasi
Department of Electronics and Communication Engineering,
RCC Institute of Information Technology, Kolkata, India
e-mail: romidey070@gmail.com

M. Banerjee
e-mail: meenakshiriabanerjee@gmail.com

A. Das
e-mail: dasantara21@gmail.com

A. Deyasi
e-mail: deyasi_arpan@yahoo.co.in

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_39

different layers [6, 7], angle of incidence [8], defect density of the materials [9], and filling factor [10]. These novel structures already proved their candidature as replacement of existing counterparts for optical communication [11, 12], opto-electronic transmitter [13], receiver [14], sensor [15], quantum computation [16], integrated optics [17], etc. It may be used as building block of nest-day communication technology.

Researches are already carried out on binary photonic crystal and its possible implementation for communication systems and devices [18, 19]. But little is carried out on ternary structures, probably due to the complexity in mathematical analysis and fabrication difficulty. Banerjee [20] showed that efficient refractometric sensing element can be designed using ternary crystal than binary one. Aly et al. [21] calculated the transmission spectra of metallic photonic crystal, and compared the result with that obtained for dielectric photonic crystal. In this paper, they have considered, binary PhC, but similar computation for ternary PhC is yet to available in literatures. Zare and Gharaati [22] calculated the bandgap width considering left-handed layer. This work is extended by Gharaati et al. [23] for 2D PhC. Sharma et al. [24] showed the design of omnidirectional filter using semiconductor materials for TPhC.

In the present paper, optical bandwidth of ternary photonic crystal (SiO_2/air/TiO_2) is computed as a function of layer dimensions and angle of incidence for both the polarization conditions. Transmittivity is calculated to study the effect of incidence angle, and filter is designed at 1550 nm for the purpose of optical communication. Result will play key factor in designing optical filter with predetermined bandwidth.

2 Mathematical Modeling

Considering the phase factor of the field propagating through uniform medium, propagation matrix is given as the function of barrier and well widths

$$P_{1,2} = \begin{pmatrix} \exp[jk_{1,2}d_{1,2}] & 0 \\ 0 & -\exp[jk_{1,2}d_{1,2}] \end{pmatrix} \tag{1}$$

where $d_{1,2}$ is the dimension of barrier/well layer, $k_{1,2}$ is the propagation vector.

Thus, transfer matrix for the elementary cell (constituting of one barrier and one well layer) is

$$M = M_1^T P_1 M_2^T P_2 \tag{2}$$

where M is the transfer matrix between the adjacent layers, given by

$$M_{1,2}^T = \frac{1}{t}\begin{pmatrix} 1 & r_{21,12} \\ r_{21,12} & 1 \end{pmatrix} \tag{3}$$

For p-polarized incident wave at angle θ_1, interface reflectivities are given by

$$r_{12} = -r_{21} = \frac{n_1 \cos(\theta_z) - n_z \cos(\theta_1)}{n_1 \cos(\theta_z) + n_z \cos(\theta_1)} \tag{4}$$

For s-polarized incident wave at angle θ_1, interface reflectivities are given by

$$r_{12} = -r_{21} = \frac{n_1 \cos(\theta_z) - n_z \cos(\theta_1)}{n_1 \cos(\theta_z) + n_z \cos(\theta_1)} \tag{5}$$

For a perfectly periodic medium composed of N such elementary cells, the total transfer matrix for such a structure is

$$M_{tot} = M_N \tag{6}$$

Transmission coefficient is given by

$$T = \frac{1}{M_{11}^2(tot)} \tag{7}$$

3 Results and Discussions

Using Eq. (7), first we have calculated the transmittivity of the ternary photonic crystal for different polarization conditions. Results are also compared with that obtained under normal incidence. Figure 1 shows the profile under s-polarized (TM mode) conditions, whereas Fig. 2 represents for p-polarized (TE mode) condition. It may be noted in this context that for both the cases, structural parameters are kept same.

From Fig. 1, it is seen that with increase of angle of incidence, notch of the transmittivity increases, which ensures about better rejection of noise. But the notch lengths on either side of desired range become asymmetric, which speaks the weakness about noise filtering. However, higher angle of incidence makes presence of less amount of ripple in passband, which is the strong point in favor of polarized incidence. Also with increase of angle, it may be observed that spectrum is shifted toward lower wavelength region, which means blueshift is taking place. In Fig. 2 (for p-polarized condition), the situation is quite different. In this case, it is found that with increasing the incidence angle, the spectrum makes right shift (redshift), and the notch of the transmittivity decreases. However, the notch length is quite higher than that observed for s-polarized condition. This helps to conclude that TE mode propagation gives better response for varying incidence angle than that obtained for TM mode propagation. In this case, less amount of ripple is observed for 5° angle of incidence, i.e., with increase of incidence angle, passband ripple varies in a zigzag fashion, not in monotonic manner.

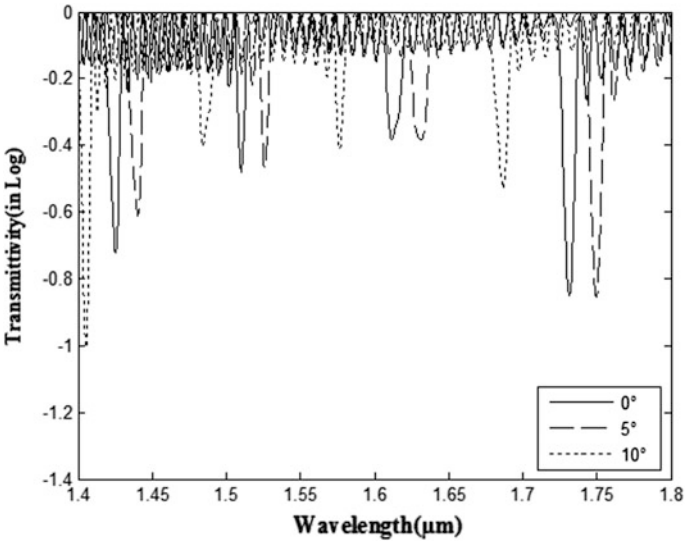


Fig. 1 Transmittivity of ternary photonic crystal under s-polarized condition and normal incidence around 1.55 μm

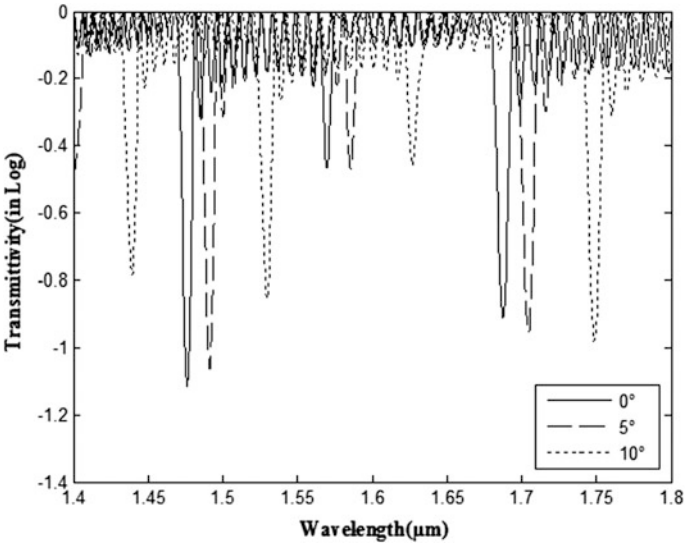


Fig. 2 Transmittivity of ternary photonic crystal under p-polarized condition and normal incidence around 1.55 μm

Figures 3 and 4 show the optical bandwidth as a function of SiO_2 layer thickness for p-polarization and s-polarization conditions, respectively. Results are compared with that obtained for normal incidence. In both the cases, bandwidth monotonically decreases with increasing SiO_2 layer width.

From Fig. 3, it may be noted that at lower dimension, bandwidth is higher for larger angle of incidence, but situation is reversed when layer thickness becomes large. The change of bandwidth is not smooth when s-polarization condition is considered. This is shown in Fig. 4. Bandwidth becomes quite large for normal incidence w.r.t angular incidence of light at this condition. A comparative study between Figs. 3 and 4 suggests that magnitude of bandwidth is smaller for s-polarized incidence than that obtained for p-polarized incidence.

Figures 5 and 6 show the optical bandwidth for varying TiO_2 dimension under both the polarization conditions. Figure 5 shows the variation for p-polarization, whereas Fig. 6 reveals the variation under s-polarization condition. Zigzag variation of bandwidth is observed for higher angle of incidence under TM mode of propagation, which is not so significant for TE mode, though a fluctuation is present. However, a comparative study between Figs. 4 and 6 reveals that higher bandwidth is obtained when SiO_2 layer is varied keeping TiO_2 layer thickness constant. This speaks in favor of filter design with required passband width, when structural parameters can be set prior to fabrication. Also passband can be tuned by changing the angle of incidence, both in magnitude as well as position.

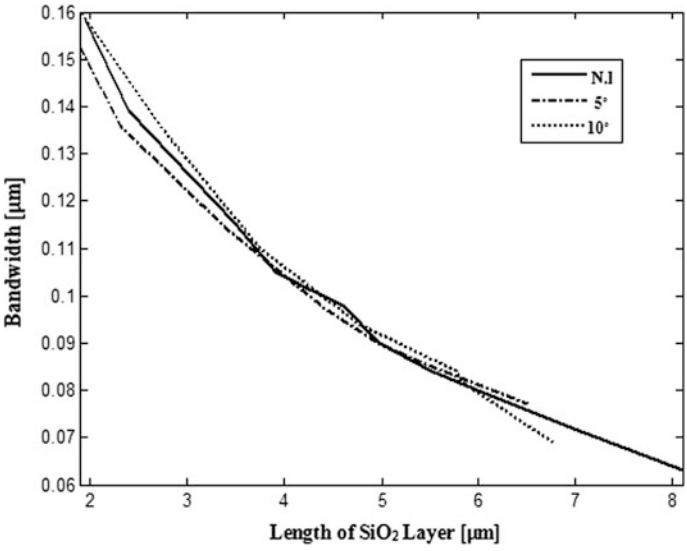


Fig. 3 Optical bandwidth as a function of SiO_2 layer thickness under p-polarized condition

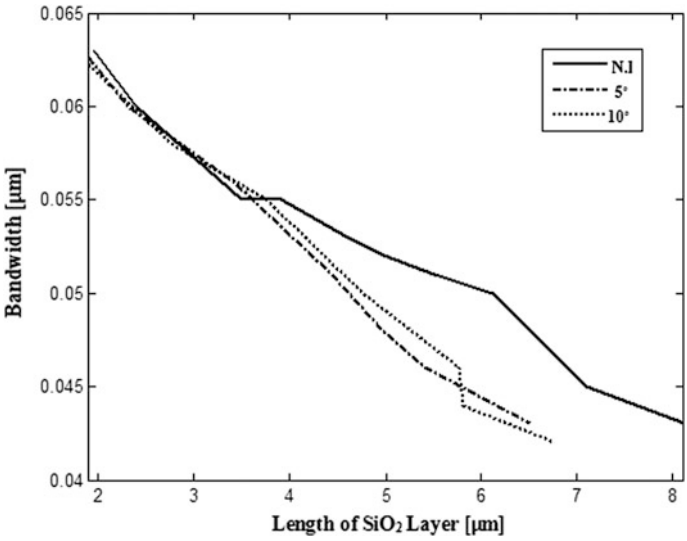


Fig. 4 Optical bandwidth as a function of SiO_2 layer thickness under s-polarized condition

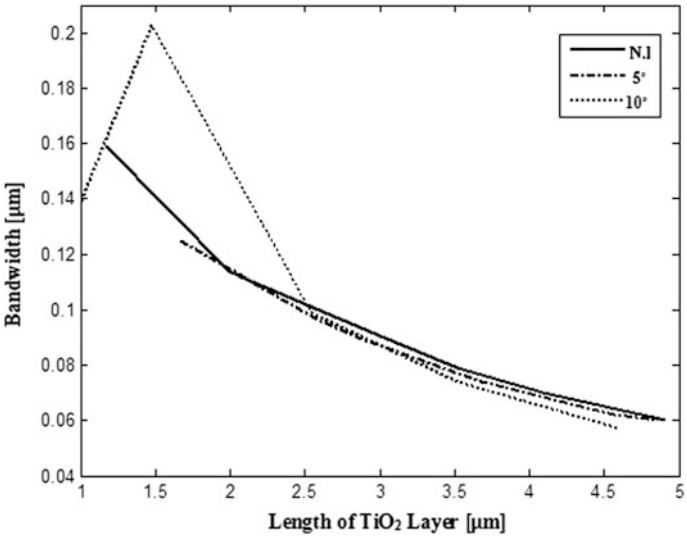


Fig. 5 Optical bandwidth as a function of TiO_2 layer thickness under p-polarized condition

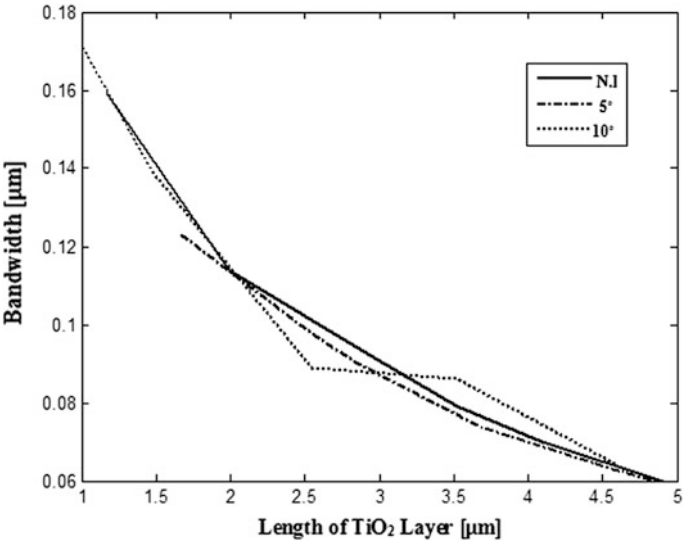


Fig. 6 Optical bandwidth as a function of TiO_2 layer thickness under s-polarized condition

4 Conclusion

Optical bandwidth and transmittivity of ternary photonic crystal are computed as a function of angle of incidence and structural parameters. Results suggest that increasing angle of incidence makes blueshift for TM mode propagation, whereas redshift is observed when TE mode is propagated. Also amount of ripple in passband depends on angle of incidence, which is a key factor in filter design. Bandwidth is decreased monotonically in general, with increasing layer thickness, whereas some zigzag variation is observed when dimension of TiO_2 layer is varied. Results have influence in designing optical filter with Butterworth characteristics in all-optical circuit.

References

1. Loudon, R.: The propagation of electromagnetic energy through an absorbing dielectric. J. Phys. A **3**, 233–245 (1970)
2. Yablonovitch, E.: Inhibited spontaneous emission in solid-state physics and electronics. Phys. Rev. Lett. **58**, 2059–2061 (1987)
3. Andreani, L.C., Agio, M., Bajoni, D., Belotti, M., Galli, M., Guizzetti, G., Malvezzi, A.M., Marabelli, F., Patrini, M., Vecchi, G.: Optical properties and photonic mode dispersion in two-dimensional and waveguide-embedded photonic crystals. Synth. Metals **139**, 695–700 (2003)

4. Villa-Villa, F., Gaspar-Armenta, J.A., Mendoza-Su´arez, A.: Surface modes in one dimensional photonic crystals that include left handed materials. J. Electromag. Waves Appl. **21**, 485–499 (2007)
5. Edalati, A., Boutayeb, H., Denidni, T.A.: Band structure analysis of reconfigurable metallic crystals: effect of active elements. J. Electromag. Waves Appl. **21**, 2421–2430 (2007)
6. Chen, J.C., Haus, H.A., Fan, S., Villeneuve, P.R., Joannopoulos, J.D.: Optical filters from photonic band gap air bridges. J. Lightwave Technol. **14**, 2575–2580 (1996)
7. Golmohammadi, S., Moravvej-Farshi, M.K., Rostami, A., Zarifkar, A.: Spectral analysis of the Fibonacci-class one-dimensional quasi-periodic structures. Progr. Electromag. Res. **75**, 69–84 (2007)
8. Mao, D., Ouyang, Z., Wang, J.C.: A photonic-crystal polarizer integrated with the functions of narrow bandpass and narrow transmission angle filtering. Appl. Phys. B **90**, 127–131 (2008)
9. Rojas, J.A.M., Alpuente, J., L´opez-Esp´i, P., Garc´ia, P.: Accurate model of electromagnetic wave propagation in unidimensional photonic crystals with defects. J. Electromag. Waves Appl. **21**, 1037–1051 (2007)
10. Mukherjee, S., Roy, A., Deyasi, A., Ghosal, S.: Dependence of photonic bandgap on material composition for two-dimensional photonic crystal with triangular geometry. Found. Front. Comput. Commun. Electr. Eng. (CRC Press), chapter 52, 259–263 (2016)
11. Limpert, J., Liem, A., Reich, M., Schreiber, T., Nolte, S., Zellmer, H., Tünnermann, A., Broeng, J., Petersson, A., Jakobsen, C.: Low-nonlinearity single-transverse-mode ytterbium-doped photonic crystal fiber amplifier. Opt. Express **12**, 1313–1319 (2004)
12. Hansryd, J., Andrekson, P.A., Westlund, M., Li, J., Hedekvist, P.O.: Fiber-based optical parametric amplifiers and their applications. IEEE J. Sel. Topics Quantum Electr. **8**, 506–520 (2002)
13. D'Orazio, A., De Palo, V., De Sario, M., Petruzzelli, V., Prudenzano, F.: Finite difference time domain modeling of light amplification in active photonic bandgap structures. Progr. Electromag. Res. **39**, 299–339 (2003)
14. Kalchmair, S., Detz, H., Cole, G.D., Andrews, A.M., Klang, P., Nobile, M., Gansch, R., Ostermaier, C., Schrenk, W., Strasser, G.: Photonic crystal slab quantum well infrared photodetector. Appl. Phys. Lett. **98**, 011105 (2011)
15. Belhadj, W., AbdelMalek, F., Bouchriha, H.: Characterization and study of photonic crystal fibres with bends. Mater. Sci. Eng.: C **26**, 578–579 (2006)
16. Azuma, H.: Quantum computation with kerr-nonlinear photonic crystals. J. Phys. D: Appl. Phys. **41**, 025102 (2008)
17. Bayat, G., Rafi, G.Z., Shaker, G.S.A., Ranjkesh, N., Chaudhuri, S.K., Safavi-Naeini, S.: Photonic-crystal based polarization converter for terahertz integrated circuit. IEEE Trans. Microwave Theory Tech. **58**, 1976–1984 (2010)
18. Gao, Y., Chen, H., Qiu, H., Lu, Q., Huang, C.: Transmission spectra characteristics of 1D photonic crystals with complex dielectric constant. Rare Metals **30**, 150–154 (2011)
19. Reininger, P., Kalchmair, S., Gansch, R., Andrews, A.M., Detz, H., Zederbauer, T., Ahn, S.I., Schrenk, W., Strasser, G.: Optimized photonic crystal design for quantum well infrared photodetectors. Proc. SPIE **8425**, 84250A (2012)
20. Banerjee, A.: Enhanced refractometric optical sensing by using one-dimensional ternary photonic crystals. Progr. Electromagn. Res. **89**, 11–22 (2009)
21. Aly, A.H., Ismaeel, M., Abdel-Rahman, E.: Comparative study of the one dimensional dielectric and metallic photonic crystals. Opt. Photonics J. **2**, 105–112 (2012)
22. Zare, Z., Gharaati, A.: Investigation of band gap width in ternary 1d photonic crystal with left-handed layer. ACTA Physica Polonica A **125**, 36–38 (2014)
23. Gharaati, A., Mohamadebrahimi, L., Roozitalab, Z.: Photonic band gap in negative ternary refractive indices of two-dimensional photonic crystal. Optica Applicata **XLIV**, 637–648 (2014)
24. Sharma, S., Kumar, R., Singh, K.S., Jain, D.: Design of an omnidirectional reflector using one dimensional ternary photonic crystal. Int. J. Eng. Tech. Res. 90–93 (2014)

Energy-Efficient Connected Target Coverage in Multi-hop Wireless Sensor Networks

Swagata Biswas, Ria Das and Punyasha Chatterjee

Abstract Wireless sensor networks (WSNs) employ numerous sensor nodes possessing sensing, processing, and wireless communication abilities to monitor a specified sensing field. As sensor nodes are mostly battery operated and are highly constrained regarding energy resources, it is essential to explore energy optimization methods to prolong WSN lifetime. Target tracking is a very conventional WSN application that demands both useful and coherent energy management. This paper proposes a distributed shortest path data collection algorithm for connected target coverage to maximize WSN lifetime pertaining to both static and mobile multi-hop WSNs. The performance is evaluated in TinyOS employing the TOSSIM simulator based on the parameters like percentage of alive nodes, load distribution of nodes, and network lifetime.

Keywords Wireless sensor network · Target coverage · Energy-efficient · Network lifetime

1 Introduction

A wireless sensor network (WSN) comprises multiple sensor nodes, which are deployed over specified sensing field to sense information about the surroundings and send the data to the sink node. WSN networks have gained unprecedented attention of researchers nowadays owing to its huge volume of applications in multiple domains like military surveillance, environmental monitoring, health care, forest fire detection, etc.

S. Biswas (✉) · R. Das · P. Chatterjee
School of Mobile Computing and Communication, Jadavpur University, Kolkata, India
e-mail: swagata4121990@gmail.com

R. Das
e-mail: ria.das76@gmail.com

P. Chatterjee
e-mail: punyasha.chatterjee@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_40

A critical and elementary issue in WSNs revolves around the coverage problem. Coverage problem in WSNs can be defined by the fact that how well the deployed sensors monitor or track an area of interest [1]. Broadly coverage problem is classified into three different categories: area coverage [2], target coverage [3], and barrier coverage [4]. In area coverage, the main goal is to monitor an area of interest. In target coverage, the prime objective is to track a set of monitoring targets. And in barrier coverage, the problem is to examine the maximal support/breach paths that traverse a sensor field. We are mostly concentrating on target coverage problem in WSNs in this paper. The general prerequisite of target coverage lies in the fact that each target should be traceable by at least one sensor node [5]. Also, after deployment, it is imperative that sensor nodes track all the targets continuously for maximum possible duration.

The basic functions of a standard sensor node comprise data sensing, processing, and communication abilities and each of these operation demands energy which is supported by the supplied battery attached with the nodes. Now, each sensor has constrained battery power and once the sensors are deployed, no further human intervention is feasible in major cases. When a sensor node completely depletes its energy source, it fails to operate and is termed as dead node. This ends the network lifetime thereby leaving some targets to be unattended although there may still persist some alive nodes in that network having enough energy to operate.

Additionally besides the sensor network coverage, it is equally significant for a sensor to maintain the network connectivity too. The term connectivity refers to the ability of every sensor node in WSN to reach their data to the sink [6]. Therefore, the problem of prolonging the network lifetime while tracking all the targets and maintaining connectivity too is a crucial task.

In this paper our objective is to sketch an energy-efficient distributed data gathering scheme for maximizing the network lifetime of WSNs deployed for monitoring a set of targets. We have proposed distributed data gathering algorithm for WSNs with static node deployment and also shown the implementation of that algorithm incorporating mobility into the network. The performance of the proposed algorithm is evaluated in Tiny OS [7] using the TOSSIM [8] simulator.

Rest of the paper is organized as follows: Sect. 2 puts forth the existing works related to target coverage problem. The problem formulation is given in Sect. 3. Section 4 presents models and definitions followed in this paper. The distributed data gathering algorithm is illustrated in Sect. 5. Section 6 discusses the simulation results while the conclusion is presented in Sect. 7.

2 Related Work

Several researchers have already explored the target coverage problem and have proposed several solutions to it. In [5], authors have already exploited the adaptive sensing range concept so as to minimize the energy consumption thereby prolonging the WSNs lifetime. In [6], authors have devised a centralized solution which addresses

the concern of power scarcity in WSN by scheduling the deployed nodes between active and sleep modes. In [9], authors have proposed a centralized power conserving scheduling algorithm by constructing maximum number of joint sets and then by determining the active time for each joint set. In [10], authors have introduced the *Connected Set Cover (CSC)* problem whose purpose is to find the maximum number of set covers such that every set is activated at different times to conserve energy. In [11], authors have proposed a greedy heuristic that prioritizes sensors based on their unused energy. In [12], authors have proposed the concept of directional sensors which can conserve energy. In [13], authors have collaborated mobile sensors with static sensors. In [14], authors introduced the concept of failure probability into target coverage problems to enhance system reliability. They have proposed a greedy algorithm that effectively computes the required number of *Maximal Reliable Sensor Covers (MRSC)*. In [15], authors have suggested a *Distributed Optimum Coverage Algorithm (DOCA)* where a large network is subdivided into many *h-hop* local networks.

3 Problem Formulation

We have mostly focused on the target coverage issue in this paper for a randomly deployed sensor network. Now, targets can be broadly defined in two ways: targets form a contiguous region [16] and targets as discrete points [6]. Here, we have considered the second one.

A sensor node i can sense or monitor any interested target t if the *Euclidean distance* between i and t, i.e., $d(i,t) \leq S_i$, the sensing range of the sensor node i. So, the *Target Coverage Problem* can be defined as [14]:

Definition 1 (*Target Coverage Problem*) Given a set of T targets and a WSN with n sensor nodes deployed over the monitoring field, target coverage problem is defined such that all the targets are continuously monitored or covered by at least one sensor node at any time.

Definition 2 (*Source node*) The sensors which sense the targets and initiate data packets into the network are called *source nodes* [6].

Sensed data from sensor nodes are conveyed to the sink node either by means of direct communication or through multi-hop paths via some intermediate nodes (*relay nodes*). Hence, in order to transfer the data to sink node, there must exist a path from each *source node* to the sink. So, the *Connected Target Coverage Problem* in multi-hop WSN can be defined as [6]:

Definition 3 (*Connected Target Coverage (CTC) Problem*) Given T targets and a WSN with n sensors, it is required that each target is covered by at least one sensor at any time and from each source to the sink, there must exist a multi-hop route.

In this paper, our goal is to compute energy-efficient data collection paths from each source node to the sink node such that maximum network lifetime can be exploited.

4 Definitions and Models

This section describes some definitions and models followed in this paper.

4.1 Definitions

Definition 4 (*Topology Graph*) Given a WSN with a set of n sensor nodes $V = \{1, 2, \ldots, n\}$ and a sink node, distributed over a *2-D* region, the network topology is represented by a graph $G(V, E)$, where an edge $(i, j) \in E$, $i, j \in V$, if and only if nodes i and j can communicate directly. $G(V, E)$ is defined as the *topology graph* of the given WSN.

Remark 1 The links in $G(V, E)$ are assumed to be bidirectional so that $G(V, E)$ is an undirected graph.

Definition 5 (*Hop-count*) In $G(V, E)$ the number of hops along the shortest path from any node i to the sink node is termed as the hop-count of node i, i.e., h_i.

Definition 6 (*Neighbors*) The set of nodes $X_i \subseteq V$ is called the *neighbors* of node i, if in $G(V, E)$, $\forall j \in X_i$ are adjacent to node i.

Definition 7 (*Network Lifetime*) The network lifetime is defined as the time duration from the time when the network started data gathering until one or more targets cannot be covered due to coverage or connectivity failure, or the first node dies out.

4.2 Network Model

The homogeneous sensor nodes and the discrete static targets are randomly distributed over a *2-D* monitoring area. The *sensing range (S)* of a node is twice the *transmission range (R)* of it to maintain the connectivity of the *topology graph*. A sensor node depletes energy during transmitting and receiving radio data. The energy consumed during sensing and mobility is considered negligible. Data communication links are assumed to be bidirectional and the antennas on sensor nodes are omnidirectional. Sink has unlimited amount of energy and computing capabilities. The positions of the sink and targets are fixed and known globally.

4.3 Traffic Model

This paper adopts a slightly modified version of the continuous traffic model, where instead of each node, at each round, only the *source node* generates a single data packet of fixed size per data gathering round. On the contrary, a *relay node* does not generate any packet but forwards others packets towards the sink. The sensed data from the *source nodes* reach the sink via the intermediate *relay nodes*.

5 Distributed Data Gathering for Connected Target Coverage

Wireless sensor network is mostly deployed in rugged regions, where human intervention is not feasible always. In this section, we propose a distributed data collection algorithm for *Connected Target Coverage* in WSN, where packets generated at the source nodes reach the sink via shortest (in terms of *hop-count*) energy-efficient paths. We assume a node density such that each target is always covered by at least one sensor at any time and the *topology graph* must be connected.

5.1 Data Gathering Algorithm

Our proposed distributed algorithm works mainly in three consecutive phases: *Hop-count Discovery Phase*, *Neighborhood Discovery Phase* and *Data Gathering Phase*. The different phases of the algorithm are detailed as follows:

Hop-count Discovery Phase: In this phase each sensor node calculates its *hop-count*. It can be easily done by flooding *Hop Configuration Message (HCM)* from the sink node [20].

Neighborhood Discovery Phase: In this phase, each sensor node finds out its *neighbors* by broadcasting some *Hello* packet containing its node-id.

Data Gathering Phase: This phase works in rounds. In every round, each of the *source* sensors monitors the targets and generates packets. To forward the packets, node chooses a lower hop *neighbor* with maximum remaining energy. The *relay* nodes does not generate any packet but forwards the packets in the same manner. The process continues until the packets reach the sink. In this way, the information from all the targets are transferred to the sink via multi-hop paths.

If a *source* senses multiple targets, one packet is generated for each target in each round. Also if a target is covered by multiple *source* nodes, each *source* generates separate packets of their own for the same target. So, in each round, many packets containing the similar information from each target reach the sink by multiple paths or *multi-path routing*. Thus even if a packet is lost, the information still can reach the sink. This increases reliabililty of the system by allowing redundancy.

Algorithm 1 describes the *Data Gathering Phase*.

Algorithm 1: Data Gathering Phase

```
Input: h_i, X_i, Es_i: Remaining energy of node i
Output: Next forwarding node j ∈ X_i
for each node i do
    //Source nodes....
    for each target t_i covered by i do
        if h_i = 1 then
            Send COVER message to sink with {t_i, i, Es_i};
        else
            j=select_next_node(i); //Select next forwarding node
            Send COVER message to j with {t_i, i, Es_i};
    //Relay nodes....
    if i receives COVER message from k ∈ X_i then
        Add Es_k to the list NEL(i);
        //Remaining energy list of neighbors if h_i = 1 then
            Send COVER message to sink with {t_i, i, Es_i};
        else
            j=select_next_node(i); //Select next forwarding node
            Send COVER message to j with {t_i, i, Es_i};
```

Algorithm 2: select_next_node

```
Input: h_i, X_i, h_j∀j ∈ X_i, Es_j∀j ∈ X_i
Output: Next forwarding node j ∈ X_i
maxE = 0; for each node j ∈ X_i and h_j < h_i do
    if Es_j > maxE then
        maxE = Es_j;
        Selected = j;
return(Selected);
```

Example 1 Figure 1 shows the *topology graph* of a randomly deployed sensor network. Figure 2 represents two of the probable routing paths through which the data gathered by the *source nodes* are forwarded to the sink.

5.2 Incorporating Mobility

There are myriad mobility models applicable to WSNs. In this paper, we followed the *Random Waypoint Mobility Model* [17] which is an alternative to *Random Walk Model* with *spatial* dependency. In such a model, a node stays in its location for a certain period of time (*pause time*), then the node locates its next destination randomly in the deployment area.

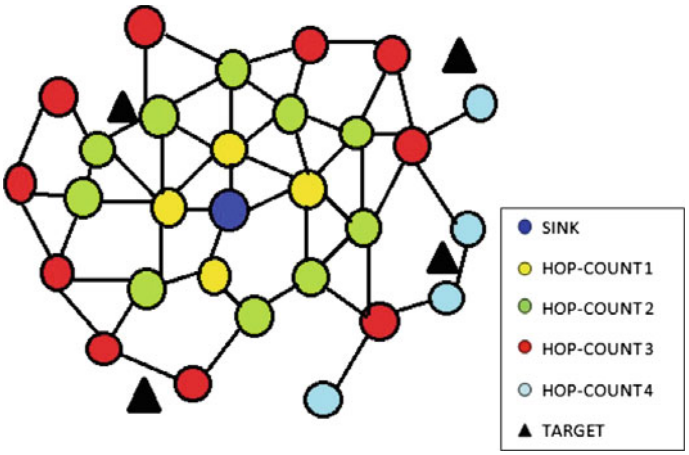


Fig. 1 Topology graph G(V, E)

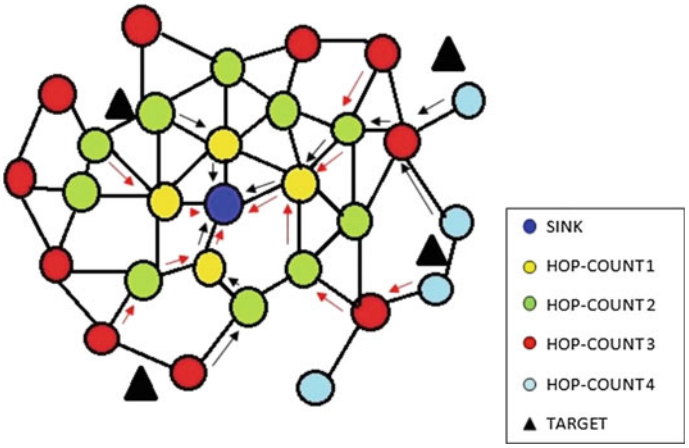


Fig. 2 Data gathering paths

In our mobility approach, we have presumed that during the *Data Gathering Phase* the sensors remain static and after some predetermined rounds of data gathering, there is a *Mobility Phase* where nodes alter their positions. During mobility the nodes do not sense or gather any information. After *Mobility Phase*, again the nodes execute *Hop-count Discovery* and *Neighborhood Discovery* phases to reconstruct the new topology.

6 Performance Evaluation

The performance of the proposed distributed algorithm is evaluated in TinyOS using the TOSSIM simulator on randomly generated connected *topology graphs*. The algorithm is developed in nesC language and python TosVis is employed for visualization purposes. The values of the default simulation parameters are presented in Table 1.

The performance of the proposed distributed data gathering algorithm is tested with respect to the following parameters: Percentage of nodes alive, load distribution of nodes and *network lifetime*.

Percentage of Nodes Alive:

In Fig. 3 we have studied the percentage of nodes alive after certain data gathering rounds for a static WSN. It is seen that about 75% of the nodes remain alive even after the *network lifetime* expires. This is because, though the *source* or *relay* nodes die out gradually due to loss of energy, there still remain other nodes in the network that neither act as a *source* nor as a *relay*. Thus their energy remains intact.

In Fig. 4, we have studied the percentage of nodes alive for a mobile WSN. Here, we infer that as compared with static WSN (Fig. 5), the network remains operational for more (almost 1.5 times) data gathering rounds.

Load Distribution of Nodes:

In Fig. 5, we have studied the load distribution of the nodes for mobile WSN when *network lifetime* expires. Load of nodes is defined as the remaining energy of nodes at any time.

Table 1 Simulation parameters

Parameter	Value
Sensing field	500×500 m^2
Sink position	Centre of the sensing field
No. of nodes deployed (n)	44
No. of targets	3
Transmission radius (T)	120 m
Sensing radius (S)	60 m
Transmission energy	1 J
Reception energy	1 J
Initial energy	500 J
Packet length	400 bits
Mobility interval	100 rounds

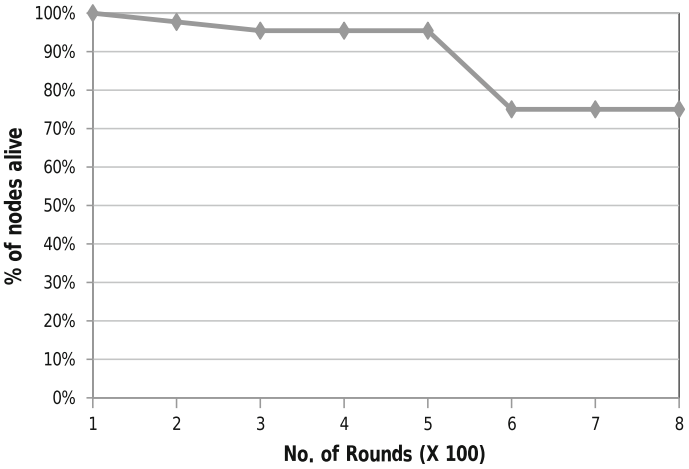


Fig. 3 Percentage of nodes alive in static WSN

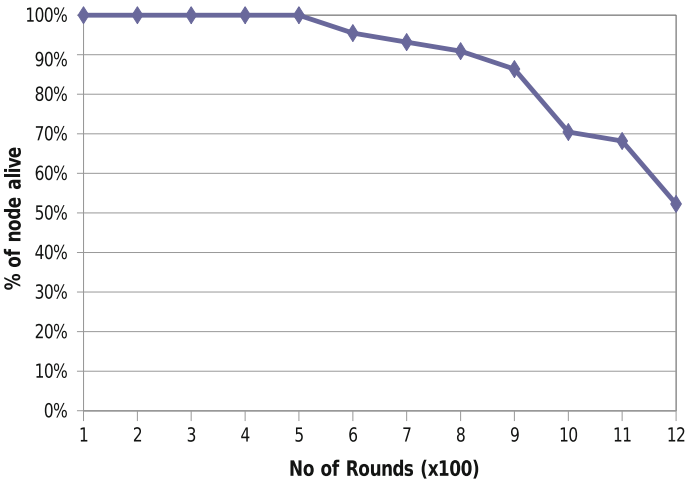


Fig. 4 Percentage of nodes alive in mobile WSN

Network Lifetime:

Table 2 gives a comparative study of the *network lifetime* for static and mobile WSNs. It is clear that the *network lifetime* is 4 times better in the first case and 2.5 times better in the second case for mobile WSN as compared with static WSN.

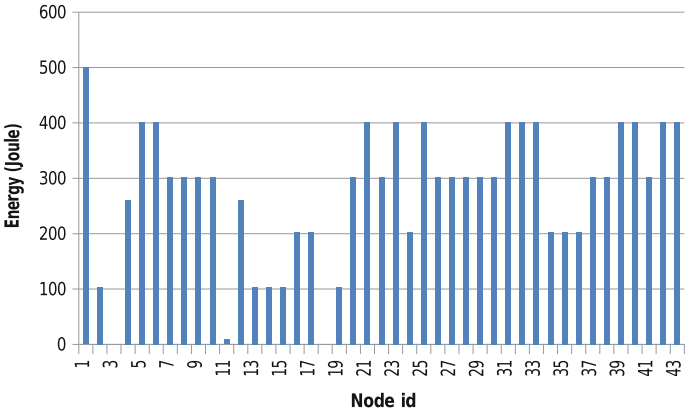


Fig. 5 Load distribution of nodes for mobile WSN

Table 2 .

Network lifetime	Static WSN	Mobile WSN
First node die out	126	506
One or more targets cannot be covered	601	1236

7 Conclusion

This paper is focused mainly on energy-efficient connected target coverage in WSN. We have proposed distributed data gathering algorithm to ensure energy-efficient shortest path routing on one hand and we have allowed multi-path routing to ensure reliability of the network on the other hand. The performance of the algorithm is analyzed on static and mobile WSNs. We observed that mobile WSN performs better both with respect to *network lifetime* and percentage of nodes alive.

Many modifications can be performed to our proposed algorithm. Data aggregation plays a significant role in saving energy. For higher degree of reliability, *k-coverage* and *k-connectivity* issues can be considered. The algorithm can be further extended to track mobile targets using other mobility models.

References

1. Huang, C.F., Tseng, Y.C.: A survey of solutions to the coverage problems in wireless sensor networks. J. Int. Technol. **6**, 1–8 (2005)
2. Carle, J., Simplot, D.: Energy efficient area monitoring by sensor networks. IEEE Comput. **37**(2), 40–46 (2004)
3. Wang, J., Niu, C., She, R.: Priority-based target coverage in directional sensor networks using a genetic algorithm. Comput. Math. Appl. **57**(11/12), 1915–1922 (2009)

4. Li, X.-Y., Wan, P.-J., Frieder, O.: Coverage in wireless adhoc sensor networks. IEEE Trans. Comput. **52**, 753–763 (2002)
5. Lu, M., et al.: Energy-efficient connected coverage of discrete targets in wireless sensor networks. ICCNMC 2005. LNCS 3619, 4352 (2005)
6. Zhao, Q., Gurusamy, M.: Lifetime maximization for connected target coverage in wireless sensor networks. IEEE/ACM Trans. Netw. **16**(6), 13781391 (2008b)
7. Farooq, M.O., Kunz, T.: Operating systems for wireless sensor networks: a survey. Sensors **11**, 5900–5930 (2011)
8. Levis, P., et al.: TOSSIM: a simulator for TinyOS networks, Version 1.0, June 26, 2003
9. Pyun, S.-Y., et al.: Power-saving scheduling for multiple-target coverage in wireless sensor networks. IEEE Commun. Lett. **13**(2) (2009)
10. Jamali, et al.: An energy-efficient algorithm for connected target coverage problem in wireless sensor networks. 978-1-4244-5540-9/10
11. Manju Pujari, A.K.: High-energy-first (HEF) heuristic for energy-efficient target coverage problem. Int. J. Ad Hoc Sens. Ubiquitous Comput. (IJASUC) **2**(1) (2011)
12. Gil, J.-M., et al.: A target coverage scheduling scheme based on genetic algorithms in directional sensor networks. Sensors **11**, 1888–1906 (2011)
13. Tan, R., et al.: Exploiting reactive mobility for collaborative target detection in wireless sensor networks. IEEE Trans. Mobile Comput. **9**(3) (2010)
14. Xiao, Y., et al.: A reliable energy efficient algorithm for target coverage in wireless sensor networks. In: 2010 IEEE 30th International Conference on Distributed Computing Systems Workshops
15. Zhang, H., Hou, J.: Maintaining sensing coverage and connectivity in large sensor networks. Ad Hoc Sens. Wireless Netw. **1**(1–2) (2005)
16. Tian, D., Georganas, N.: A coverage-preserving node scheduling scheme for large wireless sensor networks. In: Proceedings of the 1st ACM Workshop on Wireless Sensor Networks and Applications (2002)
17. Alagu Pushpa, R., et al.: Impact of mobility models on mobile sensor networks. Int. J. Commun. Netw. Secu. **1**(1) (2011)
18. Villas, L.A., et al.: DRINA: a lightweight and reliable routing approach for in-network aggregation in wireless sensor networks. IEEE Trans. Comput. **62**(4) (2013)

Optimal Sink Placement in Wireless Sensor Networks to Increase Network Performance

Mir Md. Sajid Sarwar and Punyasha Chatterjee

Abstract In traditional wireless sensor networks (WSNs), a single sink or base station is used to gather data from the whole network which suffers some serious performance issues, like latency, congestion, network failure, etc. Deploying multiple sinks in WSNs can improve the network performance, but it increases sink deployment cost. In this paper, we propose distributed algorithms to find out the minimum number of sinks to be deployed and their optimal positions in the deployment region, while ensuring a certain latency and fault tolerance level. We have considered both deterministic and random sensor node deployment strategies. Simulations of the algorithms are carried out in ns-3 network simulator, which shows number of sinks varies inversely with transmission range of nodes and network latency; and it is linearly proportional to the fault tolerance level.

Keywords Wireless sensor network · Multiple sink deployment · Hop limit · Fault tolerances · Latency · Deterministic · Random · Square grid region · Grid cross points · Optimal positions · Distributed algorithm · Network simulator ns-3

1 Introduction

A wireless sensor network (WSN) is composed of a large number of sensor nodes [1]. Each sensor node is a battery-powered device with limited storage, processing, and communication capability. It is able to sense a close-by physical phenomenon, perform a simple computation, and send its data wirelessly to a special node called

M.Md. Sajid Sarwar (✉) · P. Chatterjee
School of Mobile Computing and Communication, Jadavpur University, Kolkata, India
e-mail: sajidsarwar2011@gmail.com

P. Chatterjee
e-mail: punyasha.chatterjee@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_41

sink. The function of sink is to gather the data generated by sensors in the network. Sink normally has more capabilities than sensor nodes and it is powered by energy source.

The traditional single-sink paradigm however suffers from some major problems, as follows:

Congestion: In large wireless sensor networks, congestion into the network become very high near to the sink as all the sensor nodes send their data to the single sink. So huge number of data packets flow through the network toward the sink. This is also called *Hotspot problem* [2]. The sensors around the sink use most of its energy in relaying data packets and deplete their energy faster and die out. As a result, *energy holes* are created near to the sink which causes the sink unreachable from the rest of the sensors, even if those sensors are with sufficient residual energy.

Latency: The latency is the time duration in which a data packet reaches to the sink. More hops in routing path will directly lead to higher latency. This problem becomes worse if network size is large.

Single-point failure: In a single-sink WSN, data from the whole network is gathered at a single-sink node. If the sink fails, the whole network fails.

To overcome these issues in single-sink WSNs, multiple sinks can be deployed into the network. Deployment of multiple sinks depends on various network metrics as follows:

Acceptable latency: The number of sinks and their locations in the deployment area affects the maximum acceptable latency in the network. In real-time applications, where tolerable latency is low, sink positions should be closer to the sensor nodes to collect data at a faster rate. So, obviously more number of sinks will require.

Desired fault-tolerance level: To make a network *k-fault tolerant* [3] or *k-covered*, there should be at least k number of sinks, reachable from each sensor node.

Deployment area: If the monitoring region is large, more number of sinks will be needed to gather data from the whole network to reduce the overall latency.

Multiple sink deployment problem can be categorized into various sub-problems like sink locations/positioning problem, finding optimal number of sinks, finding efficient routes from sensors to the appropriate sinks, etc. In this paper, we have proposed distributed algorithms to find the minimum number of sinks to be deployed and their optimal positions in the deployment region, to ensure a certain latency and fault tolerance level. We have considered both deterministic and random sensor node deployment strategies.

The remainder of this paper is organized as follows: A brief account of related works is presented in Sects. 2 and 3 formulates the multiple sink deployment problem. In Sects. 4 and 5, we respectively present our approach toward for deterministic and random node deployment strategies. Section 6 presents performance evaluations and Sect. 7 concludes the paper.

2 Related Work

The deployment phase is an important aspect of wireless sensor networks. However, deployment solutions were mainly designed for sensor nodes, so as to ensure coverage and connectivity. Nevertheless, sink deployment might also be crucial as far as network lifetime, data latency, fault tolerance of the network, and various other metrics are concerned.

Finding the optimal placement of multiple sinks in a sensor network is addressed in a growing number of papers in last few years. Some of the works are discussed below.

In large-scale sensor networks, multi-sink positioning is defined as the optimization problem to maximize the performance metric like network lifetime [4], data rate or power efficiency [5–7]. The sink positioning problem is *NP-complete* even in the special case of homogeneous network [2]. In hierarchical network architecture, the complexities depend on the network clustering and sink positioning procedures [8]. Several heuristics were proposed to find near optimal solutions. Different local search techniques for sink placement in WSNs that tries to minimize the maximum worst-case delay and extend the lifetime of a WSN is presented in [3, 9]. A novel multi-sink deployment technique to optimize both the number of clusters and the cluster diameter following a graph theoretic approach is presented in [10]. In [11], authors give a mathematical model that determines the locations of the sinks minimizing the sensors' average distance from the nearest sink. In [4], authors propose multi-sink placement algorithm based on *Particle Swarm Optimization*.

3 Problem Formulation

We have considered a WSN where n homogeneous static sensors are deployed over a square monitoring region (A). Sensor nodes have transmission range T.

Then we have found the minimum number of sinks needed and their optimal locations in the deployment region such that the latency from individual node to the sink is bounded by *M-hop* and each node is reachable to k number of sinks in multi-hop routes to make the network *k-fault tolerant*. We have proposed distributed algorithms for both deterministic and random sensor nodes deployment. We also show that the minimum number of sinks needed for a network is a function: $\boldsymbol{f}(\boldsymbol{M}, \boldsymbol{K}, \boldsymbol{T})$, where M is the hop limit, k is the fault tolerance level and T is the transmission range of sensors.

To implement the distributed algorithms, we assume that the deployment region is virtually divided into square grids, where T is the diagonal of each grid block as shown in Fig. Sinks are to be placed in grid cross points. Each sink has the same transmission region T as sensor nodes.

4 Sink Placement for Deterministic Node Deployment

In deterministic node deployment strategy, sensor nodes are deployed in the monitoring region with prior planning. At the center of each grid block, a sensor node is placed.

The algorithm works in two phases as follows:

Phase 1: The sinks are placed deterministically into the network on grid cross points in such a way that there should be at least k sinks, reachable from each sensor node within its M-hop region. The algorithm (Algorithm 1) is presented below. Figure 1 shows the sink placement for *1-hop, k-covered* ($k = 1$ to 4) network.

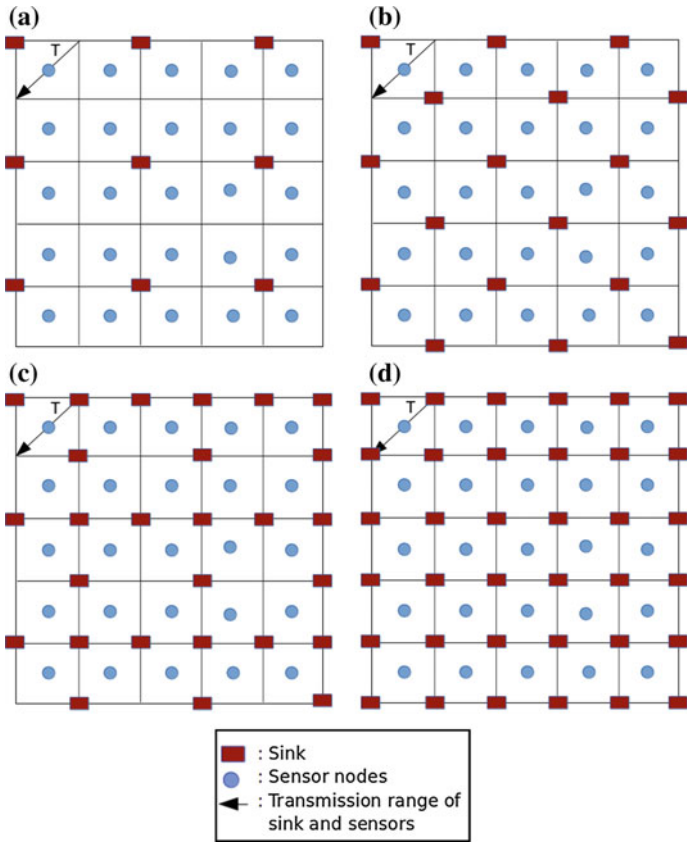


Fig. 1 Sinks placement for 1-hop: **a** 1-covered, **b** 2-covered, **c** 3-covered, **d** 4-covered network

Table 1 Variables used in Algorithm 2

Symbol	Definition
N	Set of sensor nodes
S	Set of sinks
N_jS	Set of associated sinks of node N_j

Algorithm 1: Phase 1 (Deterministic Deployment)

1. For k=1, place a sink at any corner of the grid block having diagonal M*T.
2. For k=2, place 2 sinks at any two opposite corners of the grid block having diagonal M*T.
3. For k=3, place 3 sinks at any three corners of the grid block having diagonal M*T.
4. For k=4, place sinks at all the 4 corners of the grid block having diagonal M*T

Phase 2: Each sensor node is associated with *k* number of sinks such that the latency is bounded by *M*-hop. The phase is initiated by the sinks which will broadcast a message up to *M*-hop. Sensor nodes getting those messages form their set of sinks. The algorithm (Algorithm 2) is presented below. Table 1 shows the variables used in the algorithm.

Algorithm 2: Phase 2 (Deterministic Deployment)

```
1.  For each sink Si ϵS do
2.      Send message up to M-hop
3.  end
4.  for each node Nj ϵN do
5.      if message received from Si
6.          add Si to NjS
7.      end
8.  end
```

5 Sink Placement for Random Node Deployment

In random node deployment strategy, sensor nodes are deployed in the deployment region randomly. Hence, in some grid blocks, more than one sensor node may exist, whereas some grid blocks may have no sensor nodes at all.

In our proposed algorithm, at first, we find the number and positions of candidate sinks. The candidate sink locations are found applying *Algorithm 1* on randomly deployed network, considering *1-hop, k-covered* ($k = 1$ to 4) scenario. From there we minimize it to get the optimal sink locations.

The algorithm works in three phases, as described below:

Phase 1 (Algorithm 1):

- Find the number and positions of candidate sinks.

Phase 2:

- Each sink broadcasts messages up to *M-hop*.
- Sensor nodes in the network getting these messages form a set of *Reachable Sinks* (*RS*).

Algorithm 2 can be applied here with a little modification.

Phase 3 (Algorithm 3):

- Select a node randomly from *M* alternate grid blocks as *Anchoring Node* to initiate the process.
- *Anchoring Nodes* select *k* nearest sinks (w. r. t. hop distance) from its set of *Reachable Sinks* (*RS*), and form a set of *Selected Sinks* (*SL*).
- *Anchoring nodes* broadcast *SL* up to *M-hop*.
- Other sensor nodes getting this information compare these with its own *RS*. From the sinks which are common to both *SLs* and *RS*, the node selects nearest *k* sinks as its *SL*.
- If the number of sinks in *SL* is less than *k*, it selects the rest of the sinks from its *RS* w.r.t. minimum hop distance (Fig. 2; Table 2).

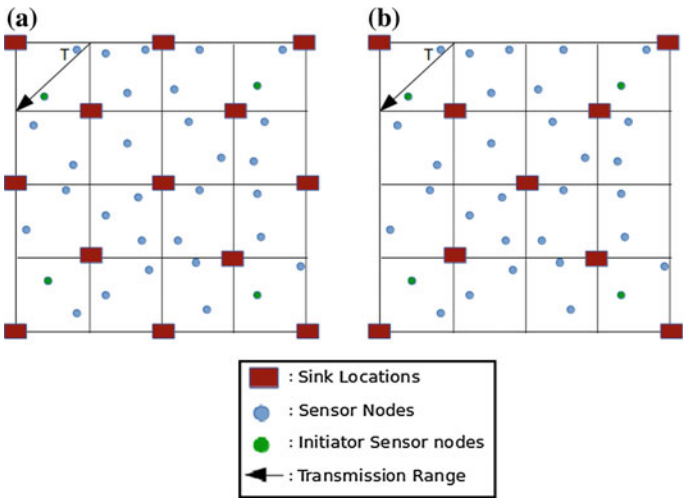


Fig. 2 **a** Candidate sink locations (2-hop, 2-covered random WSN), **b** final sink locations (2-hop, 2-covered random WSN)

Table 2 Variable Used

Symbol	Definition
N	Set of sensor nodes
S	Set of sink locations
N_iSL	Set of selected sinks of node N_i
N_jRS	Set of reachable sinks for node N_j
N_iID	Sensor node Id
N_a	Anchoring node
N_k	Sensor nodes that received message from anchoring nodes
N_KSL	Set of selected sink locations of N_k

Algorithm 3: Phase 2 (Random Deployment)

```
1.  Select a node with highest NiID from M alternate grid blocks as anchoring node
2.  for each Na do
3.      Na form set NaSL
4.      Broadcast message containing NaSL up to M-hop
5.  end
6.  for each Nk do
7.      Compare NkRS with ∑NaSL
8.      if matches then
9.          Form NkSL from ∑NaSL by selecting K sinks
10.     else
11.         Form NkSL from NkRS
12.     end
13.   end
14. end
```

Message Complexity: During the procedure, in phase 1, each sink broadcasts at most one message up to M-hop. So if there is p sinks and each sink has m number of M-hop neighbors, the number of messages will be p * n at most. In phase 2, the anchoring nodes broadcast their reachable sink sets once. This message will also propagate up to M-hop. So message complexity of this algorithm is O(m), where m is the maximum number of M-hop neighbors of a node/sink.

6 Performance Evaluation

For simulations, we have used ns-3 simulator which is a discrete event based simulator. The simulation parameters are listed in Table 3.

Table 3 Simulation parameters

OS	Ubuntu 15.04
Network Simulator	ns3
Area	500 m × 500 m
Number of Grids	5 × 5
Sensor nodes	25–50
Sensor deployment strategy	Random and deterministic
Transmission range	141 m

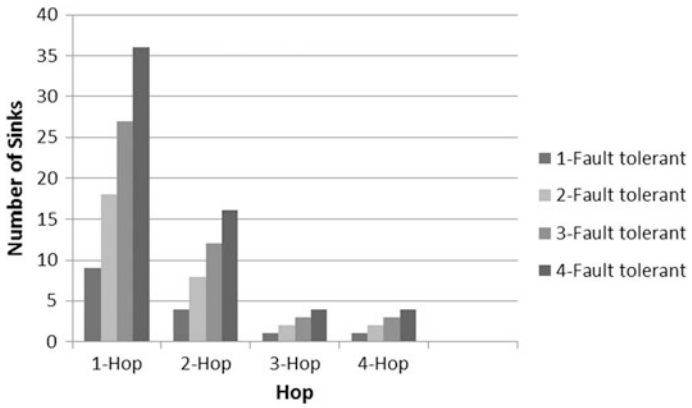


Fig. 3 Number of sinks versus hop limit

6.1 Number of Sinks in Deterministic Approach

In Fig. 3, we have shown that number of sinks needed to make a network k-covered (K = 1 to 4) while varying hop limit of the network. Number of sinks needed for the network is inversely proportional to the hop limit of the network. If we increase the acceptable hop limit or latency of the network, the number of sinks required decrease drastically.

In Fig. 4, we have shown that number of sinks needed for different hop limit and fault tolerance levels of the network while varying transmission range. It is also inversely proportional to the transmission range of the sensors.

In Fig. 5, it is shown that number of sinks needed is linearly proportional to fault tolerances level.

In wireless sensor networks, to determine the optimal number of sinks needed to meet a certain performance is very difficult task. As it depends on various metric that is discussed in earlier. From these graphs and the relationship between the number of sinks needed for a network and these parameters, we can conclude that:

Number of sinks needed for a network is a function: $f(M, K, T)$.

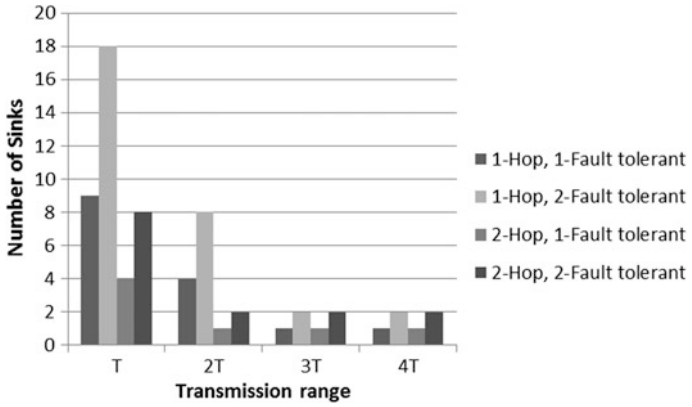


Fig. 4 Number of sinks versus transmission range

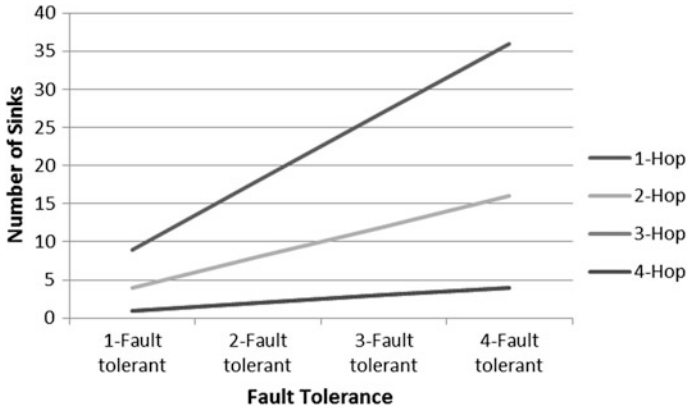


Fig. 5 Number of sinks versus fault tolerance level

6.2 Number of Sinks in Random Approach

Here, we have compared the results of sinks deployment for random sensor nodes deployment strategy with deterministic sensor nodes deployment strategy.

In Fig. 6, it is shown that in deterministic sensor node deployment for 2-hop and 2-covered wireless sensor network, number of sinks needed is 8. In simulations, for random sensor nodes deployment using our algorithm, 9 sinks get selected. So results are close to the deterministic node deployment approach.

In simulations, we only considered the case of 2-hop and 2-covered network. But it can be easily applied to the more generalize cases of M-hop and k-covered network.

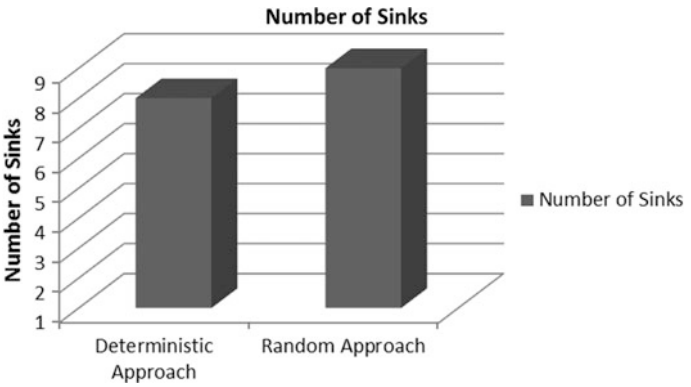


Fig. 6 Comparison between deterministic and random approach

7 Conclusions

The work carried out in this paper is focused mainly on finding the minimum number of sinks and their locations for a wireless sensor network for deterministic and random sensor node deployment strategies such that the network is k-covered and latency is bounded by M-hop. We have proposed some distributed sink placement algorithms and analyzed the performance by ns3 simulator. We found that number of sinks varies inversely with transmission range of nodes and network latency; and it is linearly proportional to the fault tolerance level. Also, our proposed algorithm, for random node deployment strategy yields very close result compared with that for deterministic node deployment.

Farther, many modifications can be done on this work. Mobility can be incorporated in the part of sensor nodes or sinks. Energy efficiency of the network can also be incorporated which will enhance the lifetime. To make the network k-fault tolerant, we have considered k = 1 to 4 only. But it can be extended further.

References

1. Akyildiz, I.F., et al.: A Survey of Sensor Networks. IEEE Commun. 102ff (2002)
2. Xu, X., Liang, W.: Placing optimal number of sinks in sensor networks for network lifetime maximization. In: IEEE International Conference on Communications (ICC) (2011)
3. Sitanayah, Lanny, et al.: Planning the deployment of multiple sinks and relays in wireless sensor networks. J. Heuristics **21**, 197–232 (2015)
4. Dandekar, D.R., Deshmukh, P.R.: Energy balancing multiple sink optimal deployment in multi-hop wireless sensor networks. In: 2013 3rd IEEE International Advance Computing Conference (IACC), pp. 408–412
5. Bogdanov, A., Maneva, E., Riesenfeld, S.: Poweraware base station positioning for sensor networks. In: Proceedings of IEEE INFOCOM 04, pp. 575–585, Hong Kong, Mar. 2004

6. Efrat, A., Har-Peled, S., Mitchell, J.S.B.: Approximation algorithms for two optimal location problems in sensor networks. In: Proceedings of 2nd International Conference Broadband Networks, vol. 1, pp. 714–723, Oct. 2005
7. Oyman, E.I., Ersoy, C.: Multiple sink network design problem in large scale wireless networks. In: IEEE International Conference on Communications (ICC), June 2004
8. Akkaya, K., Younis, M., Youssef, W.: Positioning of base stations in wireless sensor networks. IEEE Commun. Mag. **45**(4), 96–102 (2007)
9. Poe, W.Y., Schmitt, J.B.: Minimizing the maximum delay in wireless sensor networks by intelligent sink placement, Distributed Computer Systems Lab University of Kaiserslautern, 67655 Kaiserslautern, Germany, Technical Report No. 362/07
10. Chatterjee, P., Das, N.: Multiple sink deployment in multi-hop wireless sensor networks to enhance lifetime. In: Applications and Innovations in Mobile Computing (AIMoC) (2015)
11. Vincze, Z., Vida, R., Vidacs, A.: Deploying multiple sinks in multi-hop wireless sensor networks. IEEE Int. Conf. Pervasive Serv 55–63 (2007)

Performance Evaluation of Flash ADCs in Presence of Offsets Using Hot Code Generator and Bit Swap Logic (BSL)

Pranati Ghoshal and Sunit Kumar Sen

Abstract Performance of flash ADCs is beset in presence of offsets in comparators. Offsets present in comparators of a flash ADC give rise to bubble or sparkle error. There are several methods to eliminate this error—both first order or higher ones. In this paper, performance evaluation of flash ADCs will be carried out using hot code generator and bit swap logic (BSL) in presence of such offsets. It is well known that while hot code generators can take care of only first-order error in the thermometric code, BSL method can take care of any order of error. Simulation for a 3-bit flash ADC has been carried out in the presence of offsets and it has been shown that while the hot code generator can get rid of only first-order error, the BSL method overcomes any order of error.

Keywords Flash ADCs · Bit swap logic · Hot code generator · Comparator offset

1 Introduction

Flash ADCs are fastest among all the versions of ADCs available in the market. Number of comparators in a flash ADC is $2^n - 1$, where n is the number of bits in the ADC. Since the number of comparators increases exponentially in a flash ADC, it is normally contained to within 8 bits of resolution. Both power and area required for a flash ADC increase enormously as the number of bits increase. Thus a flash

P. Ghoshal (✉)
Techno India, Salt Lake, Kolkata 700091, India
e-mail: Pranati991@gmail.com

S.K. Sen
University of Calcutta, Kolkata 700073, India
e-mail: sunitksen@yahoo.co.in

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_42

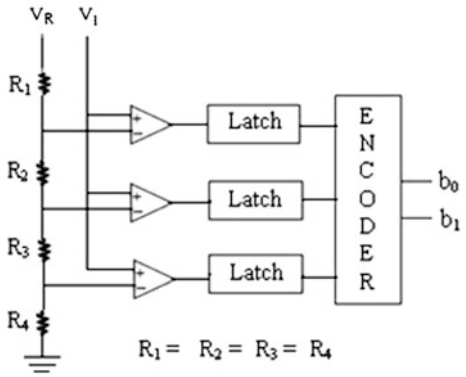


Fig. 1 Block schematic of a flash ADC

ADC is used for very fast operations, but with moderate resolution. The block diagram of a flash ADC is shown in Fig. 1. As is apparent from the figure, all the comparators operate in parallel [1]. The analog input is fed to the positive inputs of the comparators, while a predefined fixed voltage is applied at their negative inputs. The output from the comparators are in the thermometric code. In some cases, because of presence of inherent offset errors in comparators, flash ADCs suffer from bubble errors. Many researchers tried to eliminate the ragged or bubble errors present at the comparator outputs. First-order error correction schemes were implemented in [2–5].

They all use the thermometer to one-hot code (TM2OH) circuit which has the ability to correct up to first-order bubble error. A FAT tree encoder is used in [2], while ROM-based circuitries are used in [3, 4]. The FAT tree encoder offers a highly efficient encoding scheme and consumes less power compared to the ROM based circuits of [3, 4]. Mangelsdorf [5] designed a first-order bubble error correction circuitry in which the output from each comparator is compared with the outputs of its two adjacent comparators. It is changed only if it is different from either. Hieu et al. [6, 7] reported MUX-based first-order bubble error correction circuits. In [8], a suggestion was made which could correct bubble errors of any order. Disadvantages of this method were larger delay, relatively higher power consumption and a large circuit for its implementation with more output bits.

Schemes for bubble error elimination of any order was carried out in [9, 10]. In [9], Garuts et al. proposed a bit swap method which could correct bubble errors of any order, while in [10] a ROM-based scheme was implemented having NMOS transistors with pull up resistors. A bit swap logic (BSL) based bubble error correction (BEC) method was designed in [11]. It can remove bubble error of any order by bit swap logic. The degree of correction could be from first order to any higher order. Simulations in MATLAB Simulink were carried out for both for first and second orders. The designed circuitry ensured requirement of lesser number of

transistors thus leading to lesser power consumption. Kledrowetz et al. [12] analyzed the nonideal effects of comparator offsets, finite op-amp gain, and random capacitor mismatch on a pipelined ADC.

1.1 Bubble Error and Their Orders

Bubble or sparkle error at the outputs of the comparators of a flash ADC may occur due to presence of offsets present in some of the comparators. There may be variations in the error as per the offsets present and the different types of errors are enumerated as shown below:

Error types:

T_7	T_6	T_5	T_4	T_3	T_2	T_1	T_0
1st order:							
0	0	0	1	0	1	1	1
2nd order:							
0	0	1	0	0	1	1	1
3rd order:							
0	0	1	0	0	0	1	1

From the above, it is self-evident why some error is termed as first order and some other ones as second order. Depending on the severity and or particular requirements, either first- or second-order error elimination techniques are employed.

ROM-based and FAT tree encoder circuits employ thermometer to one-hot code generator (TM2OH) logic to eliminate first-order bubble error. Table 1 shows the thermometric codes, their corresponding hot codes and the equivalent binary ones. The thermometric codes shown in the table do not have any bubble error.

Table 1 Generation of thermometer to one-hot code

Thermometer code	One-hot code	Binary code
0000000	0000000	000
0000001	000000**1**	001
0000011	00000**1**0	010
0000111	0000**1**00	011
0001111	000**1**000	100
0011111	00**1**0000	101
0111111	0**1**00000	110
1111111	**1**000000	111

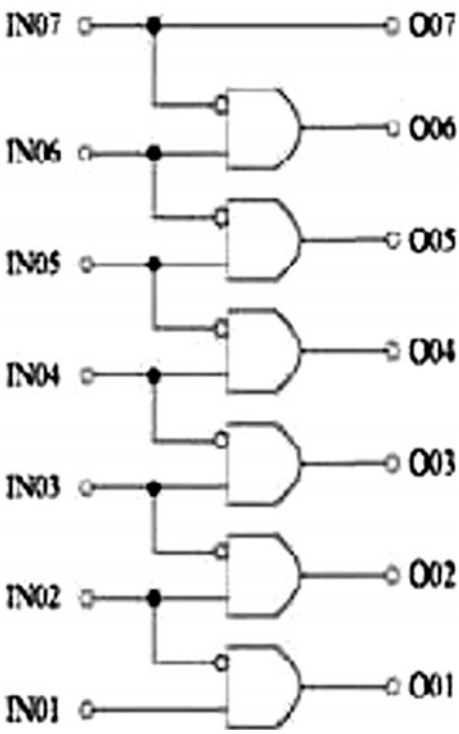


Fig. 2 TM2OH sub-circuit of a 7-to-3 ROM-based TM2B encoder

1.2 Thermometer to One-Hot Code (TMO2H) and Bit Swap Logic (BSL)

Figure 2 shows a TM2OH sub-circuit for a 7-to-3 ROM-based TM2B encoder, but it does not include any bubble error correction circuit. Thus, if a bubble error occurs anywhere during comparisons in the comparators, the output from the encoder would be erroneous. Figure 3 includes a bubble error correcting circuit which is a 7-to-3 ROM based TM2B encoder [7]. As shown in the figure, all the AND gates have three inputs with every input applied in the inverted form except the lowest input which is straightway connected.

The bit swap logic involves an error correction methodology which is capable of interchanging any out of sequence 0-1 pair—this would thus eliminate the bubble error. The methodology [9] involves bringing down a 1 and a 0 takes up its place. The process will repeat itself depending on the order of bubble error to be corrected. The bit swap logic and the corresponding truth table [10] are shown in Fig. 4.

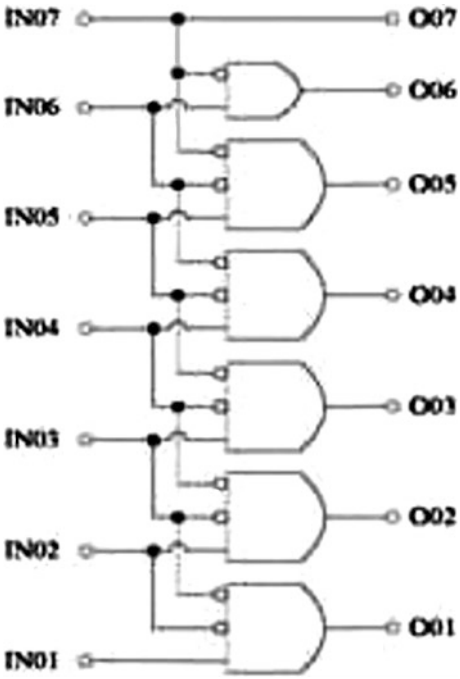
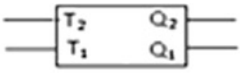


Fig. 3 TM2OH sub-circuit with BEC of a 7-to-3 ROM based TM2B encoder



T_3	T_2	Q_3	Q_2
0	0	0	0
0	1	1	0
1	0	1	0
1	1	1	1

Fig. 4 The 'Bit Swap' logic for correction of bubble error

A hardware-based BSL logic was implemented by P Ghoshal et al. [11] which is shown in Fig. 5. The input to the blocks are the ragged thermometer code and as per the order of correction needed, either Fig. 5a or 5b is used for bubble error elimination. The input to either of Fig. 5 is T_7, T_6… T_1, whereT_1 represents the lowermost comparator output, while the corrected outputs from BSL are Q_7, Q_6 … Q_1.

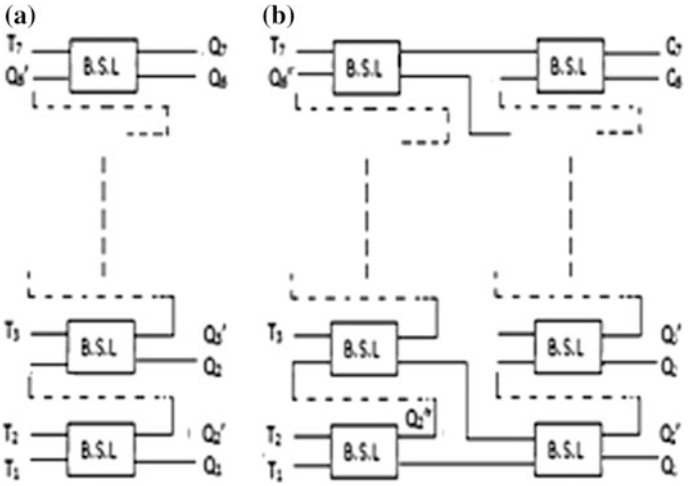


Fig. 5 Block diagram of a BSL having **a** first-order and **b** second-order bubble error correction

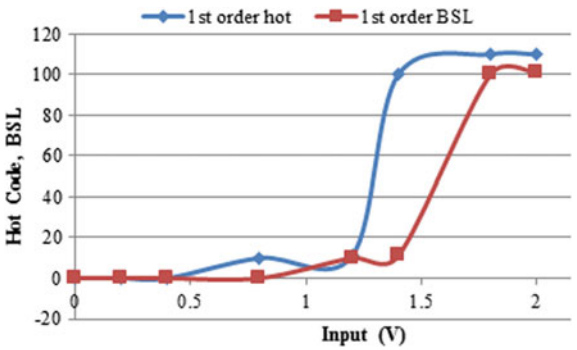


Fig. 6 Response *curves* for hot code and BSL both for first-order bubble error

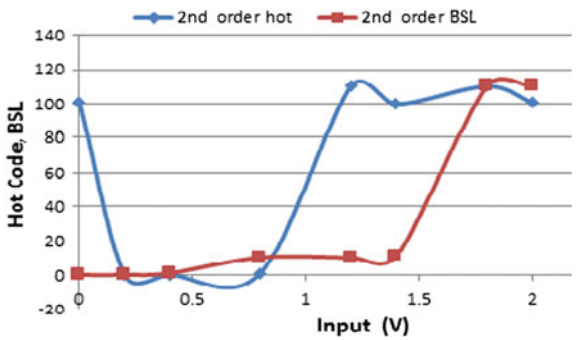


Fig. 7 Response *curves* for hot code and BSL both for second-order bubble error

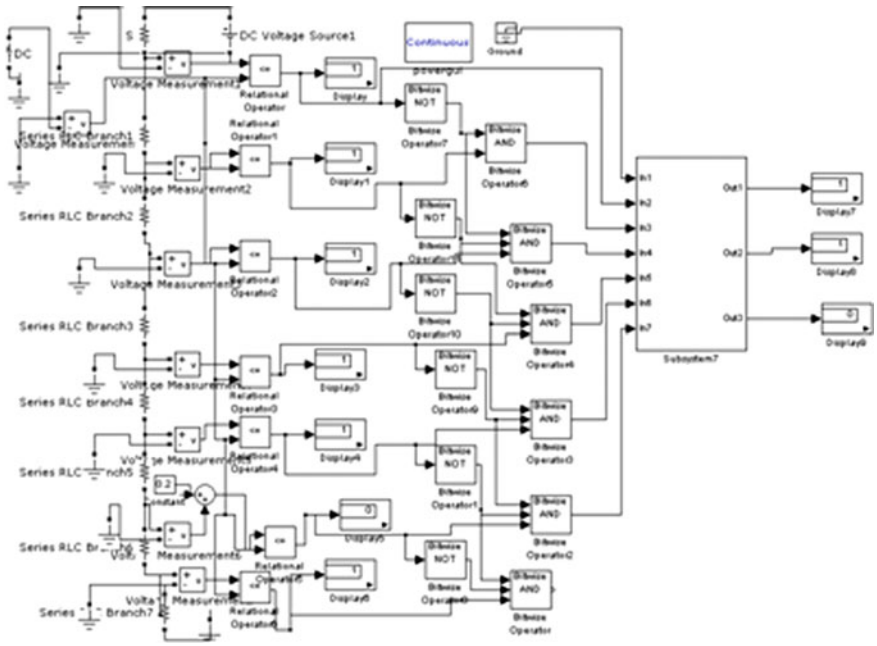


Fig. 8 Simulation for a 3-bit flash ADC with hot code generator (first-order case)

2 Implementation

Implementation will be carried out for a 3-bit flash ADC by inserting offset voltage at certain negative input(s) of the comparators—for both hot code generator circuit and BSL method. An offset voltage of value 0.2 V is applied at the negative input of comparator 2 where comparator 1 is the bottom most comparator. This is applied for both hot code and BSL methods. A reference voltage of 2 V is applied for the purpose of simulation which was done in Simulink. Response curves for hot code and BSL for first and second-order bubble error are shown in Figs. 6 and 7 respectively.

Figures 8 and 9 show the simulation diagrams for hot code circuit and BSL methods, while Table 2 shows the results for different input voltages.

In the case of second-order bubble error, offset voltages of values 0.1 and 0.3 V are applied at the negative inputs of comparators 2 and 3, respectively, for both hot code and BSL case of a 3-bit flash ADC. These are shown in Figs. 10 and 11 respectively, while the results of simulations are shown in Table 3.

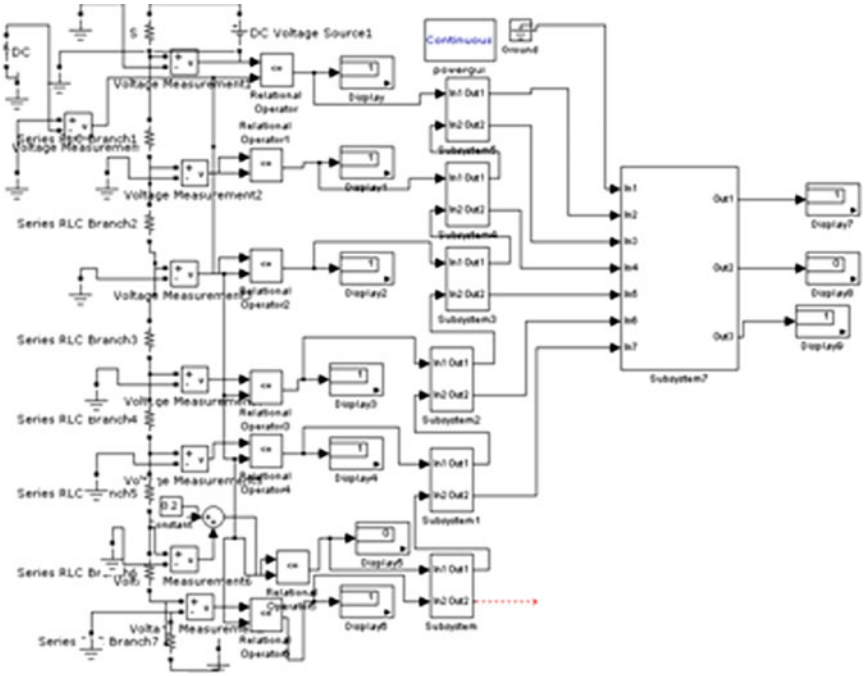


Fig. 9 Simulation for a 3-bit flash ADC with bit swap logic (first-order case)

Table 2 Simulation results for hot code and BSL for $V_{ref} = 2$ V

LP(Y)	First-order hot code	First-order BSL
0	000	000
0.2	000	000
0.4	000	000
0.8	010	000
1.2	011	010
1.4	100	011
1.8	110	100
2	110	101

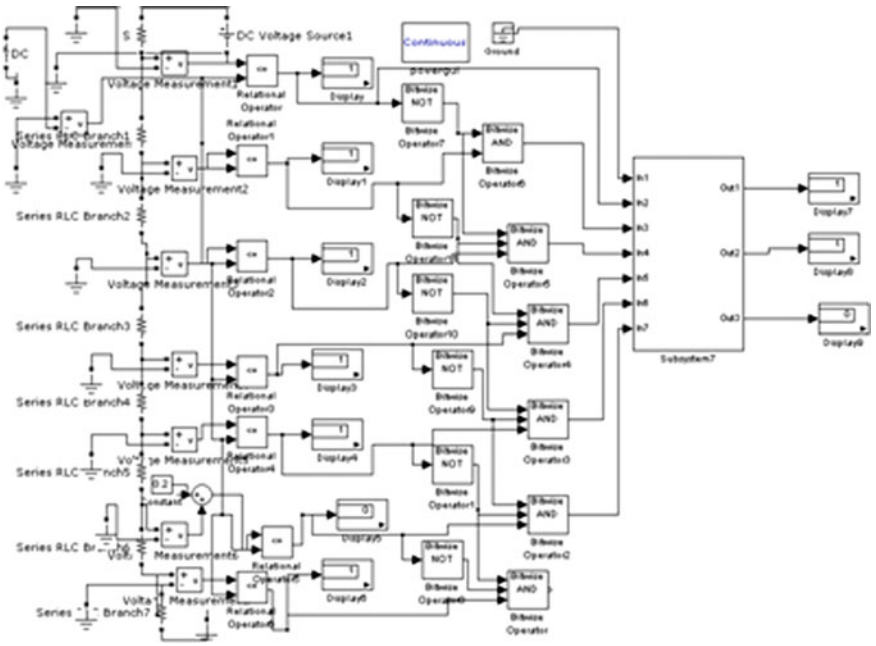


Fig. 10 Simulation for a 3-bit flash ADC with hot code generator (second-order case)

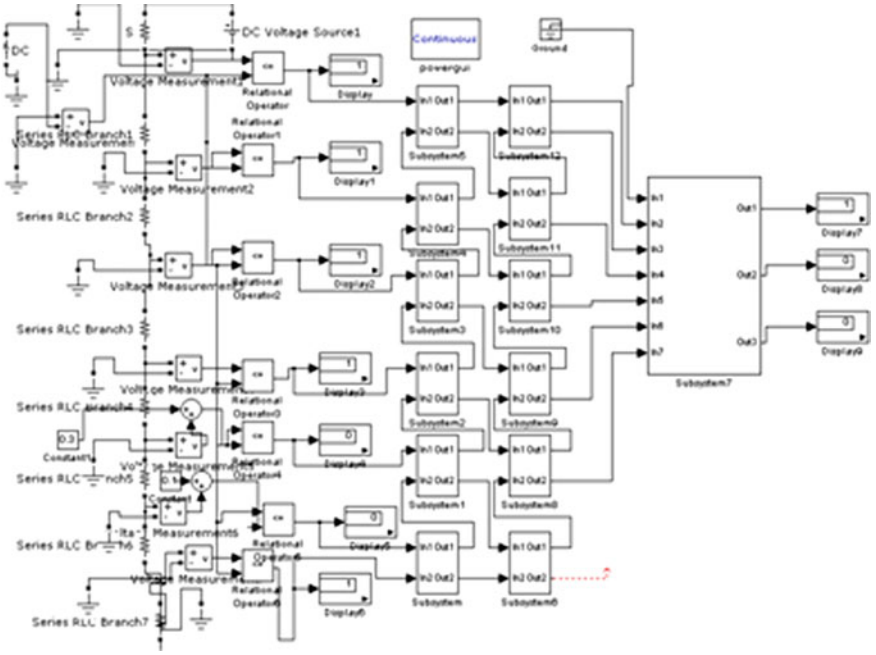


Fig. 11 Simulation for a 3-bit flash ADC with BSL (second-order case)

Table 3 Simulation results for hot code and BSL for $V_{ref} = 2$ V

I P (V)	Second-order hot code	Second-order BSL
0	101	000
0.2	000	000
0.4	000	001
0.8	001	010
1.2	110	010
1.4	100	011
1.8	110	110
2	101	110

3 Conclusion

Performance evaluation of a 3-bit flash ADC in presence of offset(s) using hot code and bit swap logic method has been reported in this paper. Simulations have been carried out for both first- and second-order bubble errors employing both the methods. The curves for either for the respective cases are shown in Figs. 6 and 7, respectively.

References

1. Kharate, G.K.: Digital Electronics. Oxford University Press (2012)
2. Daegyu, L., Cheol, Y.J., Kyusun, C., Ghaznavi, J.: Fat tree encoder design for ultra-high speed flash A/D converters, circuits and systems. In: the 45th Midwest Symposium, pp. 87–90 (2002)
3. Agarwal, N., Paily, R.: An improved ROM architecture for bubble error suppression in high speed flash ADCs. In: Proceedings of AISPC, pp. 1–5 (2008)
4. Uyttenhove, K., Marques, A., Steyaert, M.: A 6-bit 1 GHz acquisition speed CMOS flash ADC with digital error correction. In: Proceedings of the IEEE Conference, pp. 249–252 (2000)
5. Mangelsdorf, C.: A 400-MHz input flash converter with error correction. IEEE J. Solid-State Circuits **25**(1), 184–191 (1990)
6. Hieu, B.V., Choi, S., Seon, J., Oh, Y., Park, C., Park, J., Kim, H., Jeong, T.: A new approach to thermometer to binary encoder of flash ADCS—bubble error detection circuits. IEEE MWSCAS 5–8 (2011)
7. Hieu, B.V., Beak, S., Choi, S., Seon, J., Jeong, T.T.: Thermometer to binary encoder with bubble error correction (BEC) circuit for flash analog to digital converter (FADC). In: 3rd International Conference on Communications and Electronics (ICCE), pp. 102–106 (2010)
8. Wallace, C.S.: A suggestion for a fast multiplier. IEEE Trans. Electron. Comput. **13**(1), 14–17 (1964)
9. Garuts, V.E., Simon Yu, Y-C., Traa, E.O., Yamaguchi, T.: A Dual 4-bit 2-Gs/S full nyquist analog-to-digital converterusing a 70-ps silicon bipolar technology with borosenic-poly process and coupling-base implant. IEEE J. Solid-State Circuits **24** (2), 216-222 (1989)
10. Pereira, P., Fernande, J.R.: Comparative study of encoders for parallel type ADCS. In: 6th Euro Workshop on ADC Modelling and Testing, pp. 142–146 (2001)

11. Ghoshal, P., Sen, S.K.: A bit swap logic (BSL) based bubble error correction (BEC) method for flash ADCs. In: International Conference on Control, Instrumentation, Energy and Communication, CIEC-16, Jan 30–Feb 01, pp. 111–115 (2016)
12. Analysis of Non-ideal Effects of Pipelined ADC by Using MATLAB–Simulink, Advances in Sensors, Signals and Materials, pp. 85–88, ISSN: 1792-6211/ISSN: 1792-6238

Part VIII
Advanced Optimization Techniques

Optimal Sensing Time Analysis for Efficient Data Transmission in Multiple Amplify-Forward Cognitive Relay Assisted Network

Sutanu Ghosh, Aditya Chaudhuri and Sayantani Ghosh

Abstract In this paper, the performance of throughput analysis is done in an energy efficient amplify and forward (AF) relay assisted cognitive radio network. An effective optimization problem is formulated to observe the effect of two relays in proposed scenario. To achieve the desired goal, a constrained optimization problem is studied under the constraints of probability of detection threshold and maximum limitation of fixed power budget. On the basis of mathematical formulation and effective searching technique, an optimal sensing time is found to achieve high throughput of the proposed system.

Keywords Cognitive radio network · Multiple AF relay · Spectrum sensing · Throughput

1 Introduction

In the recent era, spectrum dearth issues have increased rapidly due to high demand for various wireless applications. Hence there is a need for an intelligent wireless technology for optimal utilization of spectrum. Cognitive radio (CR) [1] technology has been introduced to combat such issues where unregistered secondary users (SUs) or cognitive users (CUs) are able to access the registered spectrum while maintaining the interference level for registered primary users (PUs).

S. Ghosh (✉)
Indian Institute of Engineering Science and Technology, Shibpur, Howrah, India
e-mail: sutanu99@gmail.com

A. Chaudhuri
Dr. Sudhir Chandra Sur Degree Engineering College, Kolkata, India
e-mail: in.aditya.c@gmail.com

S. Ghosh
Jadavpur University, Kolkata, India
e-mail: qwerty123.hw@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_43

The major functionality of cognitive radio technology is accomplished by spectrum sensing (SS). There are a number of variations of SS which include waveform-based sensing [2], wavelet-based sensing [3], energy detection [4], matched filtering [5], use of polyphase filter banks [6], cyclo-stationarity based sensing [7] etc. However, these techniques have various demerits and hence cannot provide desired sensing results. So, cooperative sensing [8] methodology has been introduced to provide enhanced and reliable sensing. Moreover, relay-based cooperative sensing provide better sensing performance [9].

Previously, L. Li et al. [10] investigated optimal transmission duration and power allocation along with throughput maximization. G. Ganesan et al. [11] showed that relay assisted CR networks can reduce sensing time. Huang et al. [12] examined an Amplify-And-Forward (AF) single relay network for optimal power allocation to provide enhanced secondary throughput. Our work focuses on maximization of overall secondary throughput for a double relay AF network under the constraints of sensing performance and power budget.

The rest of the paper has been divided as follows- Sect. 2 describes the system exemplary, Sect. 3 provides the mathematical analysis based on sensing time and relay amplification, Sect. 4 includes the simulation results and the conclusions are drawn in Sect. 5.

2 Description of System Exemplary

The system exemplary used for this work has been shown in Fig. 1. It consists of a primary user (PU), cognitive source (CS), cognitive destination (CD), and two cognitive relay nodes (CR1 and CR2). This system exemplary follows the basic framework of [12] except that, here a two-way relay assisted network is considered instead of a single relay. It relies on a frame-cum-frame basis operation where each frame is divided into three time slots τ, $\frac{T-\tau}{2}$, $\frac{T-\tau}{2}$ as shown in Fig. 2. The duration of entire time frame is T msec. At the beginning of each frame, the cognitive relay network (CRN) supervises the existence of PU which highlights two scenarios (i) both the cognitive relays (CR1 and CR2) receive data from CS along with the inspection for existence of PU in the present frame if PU is not found in the preceding frame (ii) no data transmission takes place in secondary network along with inspection of PU in present frame if PU exists in preceding frame.

The channel for various links used in this exemplary is considered to be Rayleigh flat fading. The coefficients of links between PU → CS, PU → CR1, PU → CD, CS → CR1 (as well as CR1 → CS), CR1 → CD and CS → CD are depicted as h_{pcs}, h_{pcr1}, h_{pcd}, h_{csr1}, h_{crd1} and h_{csd}, respectively while for PU → CR2, CS → CR2 (as well as CR2 → CS) and CR2 → CD are represented by h_{pcr2}, h_{csr2} and h_{crd2}, respectively. Their respective distances include d_{pcs}, d_{pcr1}, d_{pcd}, d_{csr1}, d_{crd1}, d_{csd}, d_{pcr2}, d_{csr2} and d_{crd2}. The fading coefficient statistics are considered to be constant for each frame [13] and are indicated as $h_{pcs} \sim \mathcal{CN}(0, d_{pcs}^{-\alpha})$, $h_{pcr1} \sim \mathcal{CN}(0, d_{pcr1}^{-\alpha})$,

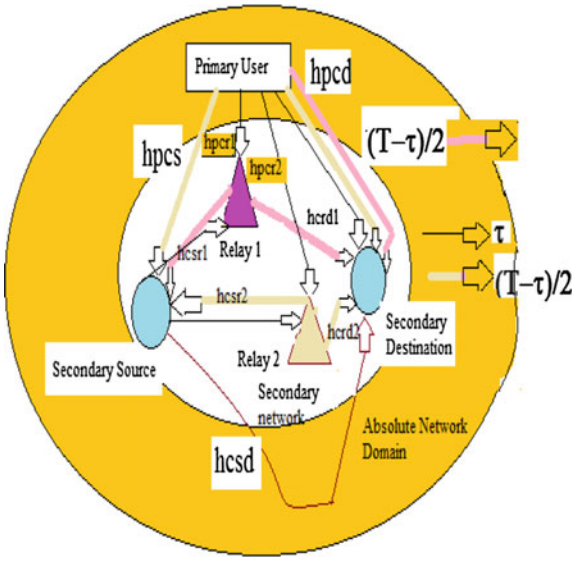


Fig. 1 System exemplary

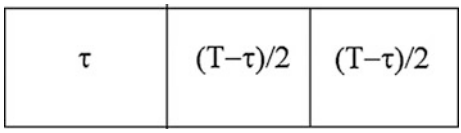


Fig. 2 Frame format for current frame M

$h_{pcd} \sim \mathcal{CN}(0, d_{pcd}^{-\alpha})$, $h_{csr1} \sim \mathcal{CN}(0, d_{csr1}^{-\alpha})$, $h_{crd1} \sim \mathcal{CN}(0, d_{crd1}^{-\alpha})$, $h_{crd1} \sim \mathcal{CN}(0, d_{crd1}^{-\alpha})$, $h_{crd1} \sim \mathcal{CN}(0, d_{crd1}^{-\alpha})$, $h_{crd1} \sim \mathcal{CN}(0, d_{crd1}^{-\alpha})$ and $h_{crd1} \sim \mathcal{CN}(0, d_{crd1}^{-\alpha})$, respectively. The path loss exponent is denoted by α. The channel gains are represented as $G_{pcs} = |h_{pcs}|^2$, $G_{pcr1} = |h_{pcr1}|^2$, $G_{pcd} = |h_{pcd}|^2$, $G_{csr1} = |h_{csr1}|^2$, $G_{crd1} = |h_{crd1}|^2$, $G_{csd} = |h_{csd}|^2$, $G_{pcr2} = |h_{pcr2}|^2$,$G_{csr2} = |h_{csr2}|^2$ and $G_{crd2} = |h_{crd2}|^2$.let $k_p(n)$ and $k_s(n)$ be the primary and secondary signal, respectively with zero mean and unit variance, i.e., $E[|k_p(n)|^2] = 1$, $E[|k_s(n)|^2] = 1$. A binary indicator Γ is used to denote the status of PU signal (Γ = 0 and 1 indicate the nonexistence and existence of PU, respectively).

At the beginning of first time frame, CR1, CR2, and CD receive primary signal as well as data from CS. So, signals obtained by CRs and CD in τcan be expressed as,

$$l_{r(i)}(n) = \sqrt{P_{ts}} h_{csr(i)} k_{s1}(n) + \vartheta \sqrt{P_{tp}} h_{pcr(i)} k_{p1}(n) + z_{r(i)}(n); \forall i \in 1, 2 \quad (1)$$

$$l_{d1}(n) = \sqrt{P_{ts}} h_{csd} k_{s1}(n) + \vartheta \sqrt{P_{tp}} h_{pcd} k_{p1}(n) + z_{d1}(n). \quad (2)$$

respectively, where n = 1, …, N, $k_{s1}(n)$ and $k_{p1}(n)$ define signals transmitted from CS and PU in τ, respectively, P_{ts} denotes transmission power of CS and P_{tp} describes transmission power of PU. N indicates the number of samples which is described by the product of τf_s, where f_s describe the sampling frequency. $z_{r(i)}(n)$ and $z_{d1}(n)$ are

the independent and identically distributed (i.i.d.), circularly symmetric complex Gaussian (CSCG) form of random sequences having zero mean and variance P_v.

During both the frames duration $\frac{T-\tau}{2}$, $\frac{T-\tau}{2}$ CR1 and CR2 amplify the signal received during τ with an amplifying gain of $\sqrt{\beta}$. Then this signal is forwarded to CS and CD. Moreover, at that instance, CS and CD receives PU signal. So, obtained signals at CS and CD can be written as: (3),

$$\begin{aligned} l_{s23}(n) &= \sqrt{\beta} h_{csr(i)} l_{r(i)}(n) + \vartheta \sqrt{P_{tp}} h_{pcs} k_{p23}(n) + z_{s23}(n) \\ &= \sqrt{\beta} \sqrt{P_{ts(i)}} h_{csr(i)} h_{csr(i)} k_{s1}(n) + \vartheta \sqrt{P_{tp}} \left(\sqrt{\beta} h_{csr(i)} h_{pcr(i)} k_{p1}(n) + h_{pcs} k_{p23}(n) \right) \\ &\quad + \sqrt{\beta} h_{csr(i)} z_{r(i)}(n) + z_{s23}(n); \quad \forall I \in 1, 2 \end{aligned} \tag{3}$$

$$\begin{aligned} l_{d23}(n) &= \sqrt{\beta} h_{crd(i)} l_{r(i)}(n) + \vartheta \sqrt{P_{tp}} h_{pcd} k_{p23}(n) + z_{d23}(n) \\ &= \sqrt{\beta} \sqrt{P_{ts(i)}} h_{crd(i)} h_{csr(i)} k_{s1}(n) + \vartheta \sqrt{P_{tp}} \left(\sqrt{\beta} h_{crd(i)} h_{pcr(i)} k_{p1}(n) + h_{pcd} k_{p23}(n) \right) \\ &\quad + \sqrt{\beta} h_{crd(i)}(n) z_{r(i)}(n) + z_{d23}(n); \forall i \in 1, 2 \end{aligned} \tag{4}$$

respectively, where n = N+1, …, 2 N, $k_{p23}(n)$ indicates the PU signal with the mean of zero and unity variance $E[|k_{p23}(n)|^2] = 1$ at τ. Noise components $z_{s23}(n)$ and $z_{d23}(n)$ are i.i.d., CSCG random sequences with zero mean and variance of P_v .

Self-interference at CS is described by the first mathematical term in (3). Hence, by applying self-interference cancelation (canceling $k_{s1}(n)$ which originates from CS) (3) can be rewritten as

$$\widetilde{l_{s23}(n)} = \vartheta \sqrt{P_{tp}} \left(\sqrt{\beta} h_{csr(i)} h_{pcr(i)} k_{p1}(n) + h_{pcs} k_{p23}(n) \right) + \sqrt{\beta} h_{csr(i)} z_{r(i)}(n) + z_{s23}(n); \forall i \in 1, 2. \tag{5}$$

The signal described in (5) assists in energy detection at CS for SS by using $\widetilde{l_{s23}(n)}$. Test statistic L_s, can be mathematically described as $L_s = \sum_{n=N+1}^{2N} \left| \widetilde{l_{s23}(n)}^2 \right|$ and follows chi-square distribution. Based on the central limit theorem, for abundant number of samples N, the distribution of test statistic K_s nearly pursues a Gaussian pattern under both the hypothesis H_0 ($\vartheta = 0$) and H_1 ($\vartheta = 1$). To simplify this, $k_{p1}(n)$ and $k_{p23}(n)$ are considered to be CSCG form of random sequences and $k_{p1}(n)$, $k_{p23}(n)$, $z_{r1}(n)$, and $z_{s23}(n)$ are pair-wise mutually independent. The mean and variance of L_s under H_1 can be written as $E(L_{s1}) = N\mu_1$ and var $(L_{s1}) = N\mu_1^2$, respectively, where $\mu_1 = G_{pcs} P_{tp} + \beta\, G_{csr(i)} (G_{pcr(i)} P_{tp} + P_v) + P_v; \forall i \in 1, 2$. Under the hypothesis H_0, L_s can be expressed as $E(L_{s0}) = N\, \mu_0$ and var(L_{s0})=N var$(L_{s0}) = N\, \mu_0^2$, respectively, where $\mu_0 = \beta\, G_{csr(i)} P_v + P_v; \forall i \in 1, 2$.

Let, χ denote the decision threshold for detection probability p_d and false alarm probability p_f. These probabilities can be expressed as follows:

$$p_d = Q\left(\frac{(\chi - \mathrm{E}(\mathrm{L_{s1}}))}{(\sqrt{\mathrm{var}(\mathrm{L_{s1}})})}\right) = Q\left(\frac{(\chi - (\mathrm{N}\ \mu_1))}{\sqrt{\mathrm{N}}\mu_1}\right). \tag{6}$$

$$p_f = \mathrm{Q}\left(\frac{(\chi - \mathrm{E}(\mathrm{L_{s0}}))}{(\sqrt{\mathrm{var}(\mathrm{L_{s0}})})}\right) = \mathrm{Q}\left(\frac{(\chi - (\mathrm{N}\ \mu_0))}{\sqrt{\mathrm{N}}\mu_0}\right). \tag{7}$$

where, Q(.) denotes the complementary cumulative distribution function of standard form of Gaussian random variable.

Now, to meet the target threshold of detection $\mathrm{p_d^{tr}}$, from (6) and (7), by canceling the decision threshold χ , false alarm probability $\mathrm{p_f}$ can be expressed in the following way:

$$\mathrm{p_f} = \mathrm{Q}\left(\frac{\mathrm{Q}^{-1}\left(\mathrm{p_d^{tr}}\right)\mu_1 + \sqrt{\mathrm{N}}(\mu_1 - \mu_0)}{\mu_0}\right) = \mathrm{Q}\left(\left(\mathrm{Q}^{-1}\left(\mathrm{p_d^{tr}}\right) + \sqrt{\mathrm{N}}\right)\gamma_{\mathrm{pp}} + \mathrm{Q}^{-1}\left(\mathrm{p_d^{tr}}\right)\right) \tag{8}$$

where, $\gamma_{\mathrm{pp}} = \frac{(\mu_1 - \mu_0)}{(\mu_0)} = \frac{\gamma_{\mathrm{pcs}} + \beta\ \mathrm{G_{csr(i)}}\gamma_{\mathrm{pcr(i)}}}{1 + \beta\ \mathrm{G_{csr(i)}}}$; $\quad \forall i \in 1, 2.$

$$\gamma_{\mathrm{pcs}} = \frac{\mathrm{G_{pcs}P_{tp}}}{\mathrm{P_v}}\ \&\ \gamma_{\mathrm{pcr(i)}} = \frac{\mathrm{G_{pcr(i)}P_{tp}}}{\mathrm{P_v}}; \quad \forall i \in 1, 2.$$

γ_{pcs} and $\gamma_{\mathrm{pcr(i)}}$ are transmitted signal SNR along the link of PU to CS and PU to both the CRs, respectively.

3 Mathematical Analysis on Sensing Time and Relay Amplification

The secondary network relay throughput can be calculated using three equations. First two equations are about the individual throughput of the two relays and the last one is the total sum of those throughputs. So, the secondary network relay throughput in CRN can be mathematically expressed as

$$R_a = \left(\frac{\mathrm{T} - \tau}{2}\right)(1 - \rho)\left(1 - p_{f1}\right)\left(log_2(1 + \gamma_{tr1})\right) \tag{9}$$

$$R_b = \left(\frac{\mathrm{T} - \tau}{2}\right)(1 - \rho)(1 - p_{f2})(log_2(1 + \gamma_{tr2})) \tag{10}$$

$$R_{total} = R_a + R_b \tag{11}$$

where, R_a, R_b are the individual throughput for relay CR1 and CR2 respectively.R_{total} is the sum of the individual throughput for CR1 and CR2.γ_{tr1} and γ_{tr2} indicates the cognitive signal-to-noise ratio (CSNR) received at CD for relay CR1 and CR2, respectively during the inactive state of PU ($\vartheta = 0$).ρ depicts probability when PU is in active state, i.e., $\rho = \mathrm{P}(\vartheta = 1) = \mathrm{P}(\mathrm{H}_1)$. CD uses maximal-ratio combining technique to calculate the throughput of CS and CRs received by CD during time frames τ , $\frac{\mathrm{T}-\tau}{2}$ and $\frac{\mathrm{T}-\tau}{2}$, respectively. So, using Eqs. (2) and (4) CSNR can be written as

$$\gamma_{tr1} = \frac{G_{csd}P_{ts1}}{\mathrm{P_v}} + \frac{\beta G_{crd1}G_{csr1}P_{ts1}}{\beta G_{crd1}\mathrm{P_v} + \mathrm{P_v}} \tag{12}$$

$$\gamma_{tr2} = \frac{G_{csd}P_{ts2}}{\mathrm{P_v}} + \frac{\beta G_{crd2}G_{csr2}P_{ts2}}{\beta G_{crd2}\mathrm{P_v} + \mathrm{P_v}} \tag{13}$$

From (10), transmission powers for the two relays are represented, respectively

$$P_{R(i)} = \left(\beta G_{csr(i)}P_{ts(i)} + \beta\vartheta G_{pcr(i)}P_{tp} + \mathrm{P_v}\beta\right); \text{for } \forall \mathrm{i} \in 1, 2 \tag{14}$$

The average transmission power for the two relays can be written as

$$\begin{aligned}\overline{P_R} &= \rho\left(\beta G_{csr(i)}P_{ts(i)} + \beta G_{pcr(i)}P_{tp} + \mathrm{P_v}\beta\right) + (1-\rho)\left(\beta G_{csr(i)}P_{ts(i)} + \mathrm{P_v}\beta\right); \quad \forall i \in 1,\ 2 \\ &= \left(\beta G_{csr(i)}P_{ts(i)} + \beta\rho G_{pcr(i)}P_{tp} + \mathrm{P_v}\beta\right);\end{aligned} \tag{15}$$

The objective of this work is to find the optimal sensing time for best performance between two relay nodes and to maximize the overall secondary throughput satisfying sensing constraint to protect PU and maintaining a total transmission power budget P_{opt} of two relay nodes.

Then optimization problem can be formulated in the following way:

Problem:

$$\max_{\tau}(R_1, R_2) \tag{16a}$$

$$\max_{\tau}(R_{total}) \tag{16b}$$

Under the constraints

$$\mathrm{c}_1 : p_d \geq p_d^{tr}$$

$$\mathrm{c}_2 : \beta\left(G_{csr(i)}P_{ts(i)} + \rho G_{pcr(i)}P_{tp} + P_v\right) \leq P_{opt};\ \forall \mathrm{i} \in 1, 2$$

$$c_3: \tau \geq 0, \frac{T-\tau}{2} \geq 0 \tag{16c}$$

where, p_d^{tr} denotes the probability detection threshold and P_{opt} indicates the total transmission power budget.

On the basis of (16b) power equality constraint can be reconstructed as

$$P_{tsi} = \frac{P_{opt} - \rho\beta G_{pcr(i)} P_{tp} - \beta P_v}{\beta G_{csr(i)}}; \forall i \in 1, 2 \tag{17}$$

Therefore by putting (8) and (17) into (16a and 16b), the constrained optimization problem is converted into unconstrained optimization problem. Optimal value of τ^* can be found using an efficient searching technique.

4 Result Analysis

The parameter settings used for simulation include $\sigma = 0.3, p_d^{thr} = 0.95, P_v = 0$ dBW, $P_{opt} = 10$ dBW, $G_{csr1} = 0.4, G_{csr2} = 0.1, G_{crd1} = 0.1, G_{crd2} = 0.4, G_{pcr1} = 0.4, G_{pcr2} = 0.4$, $G_{csd} = G_{pcs} = 0.1, \beta = 2, T = 1$ sec and $f_s = 1000$.

Figure 3. represents the comparison between two relay transmission with respect to sensing time. From both of the concave nature of throughput, it is observed that R_a has higher rate than R_b. Figure 4 shows overall secondary throughput versus sensing time. It is seen that initially throughput increases with an increasing value of sensing time and then after reaching an optimal value, it gradually falls with higher value of sensing time. This happens because of the fact that as p_f decreases,

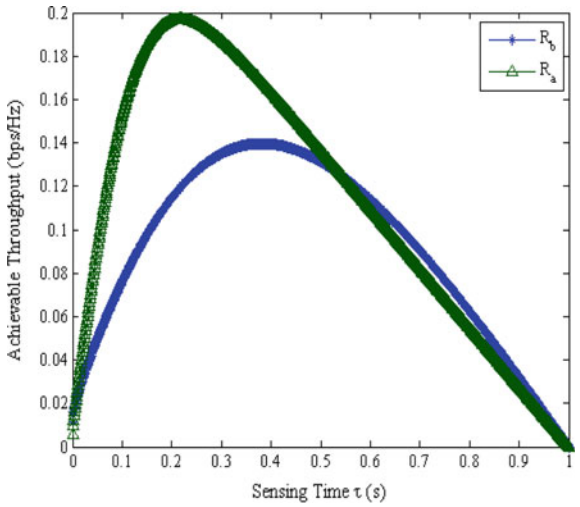


Fig. 3 Comparison of throughput with respect to sensing time for two relay transmission

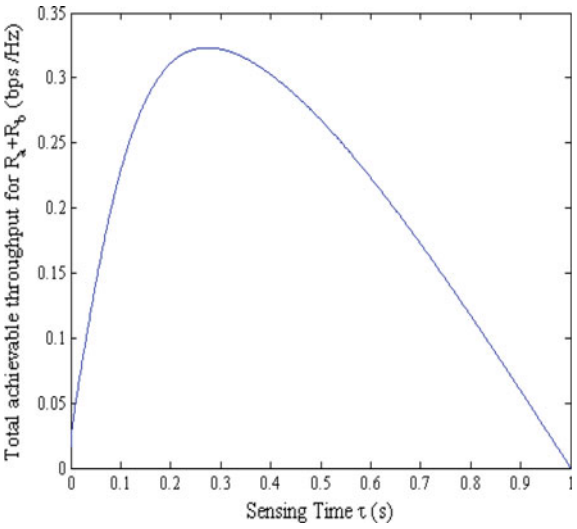


Fig. 4 Total achievable throughput versus sensing time

then secondary throughput increases due to enhancement of $(1-p_f)$ factor. However, when $(1-p_f))$ factor almost attains the peak value, then secondary throughput decreases and reduces close to zero Fig. 4.

5 Conclusion

Based on mathematical analysis and simulation results, we can conclude that an efficient throughput enhancement is obtained for two relay assisted cognitive radio network. Additionally, an optimum value of sensing time has also been found to achieve the high rate of transmission for the system. Moreover, in future this work can be extended for multiple decode and forward relay assisted cognitive radio network scenario.

References

1. Zhang, R., Liang, Y.C., Cui, S.: Dynamic resource allocation in cognitive radio networks. IEEE Signal Process. Mag. **27**(3), 102–114 (2010)
2. Tang, H.: Some physical layer issues of wide-band cognitive radio systems. In: 2005 First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks DySPAN, pp. 151–159 8 Nov 2005
3. Li, X., Hu, F., Zhang, H., Shi, C.: Two-branch wavelet denoising for accurate spectrum sensing in cognitive radios. Telecommun. Syst. **57**(1), 81–90 (2014)

4. Joshi, D.R., Popescu, D.C., Dobre, O.A.: Gradient-based threshold adaptation for energy detector in cognitive radio systems. IEEE Commun. Lett. **15**(1), 19–21 (2011)
5. Cabric, D.B.: Cognitive radios: System design perspective. ProQuest; 2007
6. Farhang-Boroujeny, B., Kempter, R.: Multicarrier communication techniques for spectrum sensing and communication in cognitive radios. IEEE Commun. Mag. **46**(4), 80–85 (2008)
7. Du, K.L., Mow, W.H.: Affordable cyclostationarity-based spectrum sensing for cognitive radio with smart antennas. IEEE Trans. Veh. Technol. **59**(4), 1877–1886 (2010)
8. Zhang, W., Mallik, R.K., Letaief, K.B.: Optimization of cooperative spectrum sensing with energy detection in cognitive radio networks. IEEE Trans. Wirel. Commun. **8**(12), 5761–5766 (2009)
9. Chen, H.: Relay selection for cooperative spectrum sensing in cognitive radio networks. In: 2010 International Conference on Communications and Mobile Computing (CMC), vol. 2, pp. 188–192, 12 Apr 2010
10. Li, L., Zhou, X., Xu, H., Li, G.Y., Wang, D., Soong, A.: Energy-efficient transmission in cognitive radio networks. In: IEEE Consumer Communications and Networking Conference (CCNC), pp. 1–5, Jan 2010
11. Ganesan, G., Li, Y.: Cooperative spectrum sensing in cognitive radio, part I: Two user networks. IEEE Trans. Wirel. Commun. **6**(6), 2204–2213 (2007)
12. Huang, S., Chen, H., Zhang, Y.: Optimal power allocation for spectrum sensing and data transmission in cognitive relay networks. IEEE Wirel. Commun. Lett. **1**(1), 26–29 (2012)
13. Li, L., Zhou, X., Xu, H., Li, G.Y., Wang, D., Soong, A.: Simplified relay selection and power allocation in cooperative cognitive radio systems. IEEE Trans. Wirel. Commun. **10**(1), 33–36 (2011)

R Implementation of Bayesian Decision Theoretic Rough Set Model for Attribute Reduction

Utpal Pal, Sharmistha Bhattacharya (Halder) and Kalyani Debnath

Abstract Bayesian Decision Theoretic Rough Set (BDTRS) model is a significant advancement in the field of attribute reduction of an information system. However, a lack of the related software for implementing this model can be observed preventing their use in practice. In this paper, the BDTRS model for attribute reduction is further studied and implemented using R programming language as functions. These new R functions are further compared with few existing sophisticated rough set based attribute reduction methods (R functions), available within "RoughSets" package of R. For comparative analysis, secondary data sets from UCI Machine Learning repository has been used. Improved results have been achieved by the implemented BDTRS-based R functions compared to other existing functions. The implemented BDTRS model may now perform attribute reduction for high-dimensional large size practical field data sets which was not possible earlier.

Keywords Attribute reduction · R implementation · Bayesian decision theoretic rough set · Comparison

1 Introduction

Ever since the introduction of rough set theory by Pawlak [7] in 1982, many extensions have been made. The results of the studies: Decision-Theoretic rough set model [14], Variable Precision rough set models by Ziarko [15], Bayesian Rough set model by Slezak and Ziarko [12], increase our understanding of the rough set

U. Pal (✉) · S. Bhattacharya (Halder) · K. Debnath
Department of Mathematics, Tripura University (A Central University), Agartala, Tripura, India
e-mail: utpalpal@tripurauniv.in

S. Bhattacharya (Halder)
e-mail: s.bhattacharya@tripurauniv.in

K. Debnath
e-mail: dkalyanimath@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_44

theory and its domain of applications [6]. A milestone is represented by BDTRS model for attribute reduction by Bhattacharya (Halder) [1, 2]. These models can be used for reduction of insignificant condition attribute from a decision table.

R [5, 8] is a widely used open source analysis environment for scientific computing and visualization. More than 8000 packages are already included in the repositories of the Comprehensive R Archive Network (CRAN) at http://cran.r-project.org. Functions written in other languages (like C++, JAVA) can easily be coupled with R functions. The R packages such as "RoughSets" [9], "stringi" are extensively used in this work. Three basic steps are followed by BDTRS model for attribute reduction. Step1, computes Bayesian Decision Theoretic Positive Region. Step2, computes Bayesian Decision Theoretic Discernibility Matrix. Step3, computes reduced attribute set, solving the Discernibility Matrix of step2.

The remaining part of this paper is organized as follows. In the next section, a brief introduction to basic (Pawlak) rough set and BDTRS is given. Section 3, in particular, discusses about processing steps, algorithms and execution process (with example) of BDTRS model in R environment. Results of comparative study are presented in Sect. 4. Finally, Sect. 5 contains concluding remarks.

2 Theoretical Background

2.1 Pawlak's Rough Set Model [7]

The upper and lower approximation of X with respect to equivalence relation R with condition probability are denoted as $\overline{R}X$ and $\underline{R}X$, respectively, and defined as follows:

$$\overline{R}X = \cup\{[x]_R | P(X/[x]_R) > 0,\ [x]_R \in \pi_R\}$$

$$\underline{R}X = \cup\{[x]_R | P(X/[x]_R) = 1,\ [x]_R \in \pi_R\}$$

Three kinds of approximation regions of X with respect to A can be defined according to its upper and lower approximation, respectively.

Positive region

$$POS_R(X) = \underline{R}X = \cup\{[x]_R | P(X/[x]_R) = 1,\ [x]_R \in \pi_R\}$$

Negative region

$$NEG_R(X) = U - \overline{R}X = \cup\{[x]_R | P(X/[x]_R) = 0,\ [x]_R \in \pi_R\}$$

Boundary region

$$BND_R(X) = \overline{R}X - \underline{R}X = \cup \{[x]_R | 0 < P(X/[x]_R) < 1,\ [x]_R \in \pi_R\}$$

2.2 Reduct and Significance of the Reducts [7]

A reduct is a minimal set of attributes from A that preserves the partitioning of the universe and hence the ability to perform classifications as the whole attribute set A does.

The consistency factor for the condition attribute C and decision attribute D is defined as

$$\gamma(C, D) = |POS_C(D)| / |U|$$

If $\gamma(\mathrm{C}, \mathrm{D}) = 1$, the decision table is consistent and if $\gamma(\mathrm{C}, \mathrm{D}) \neq 1$, the decision table is inconsistent. But the coefficient $\gamma(\mathrm{C}, \mathrm{D})$ changes when we remove any attribute $a \in \mathrm{C}$. The significance of any attributes a can be evaluated by measuring effect of removing the attribute from an information table on classification and is defined as

$$\sigma(\mathrm{C}, \mathrm{D})(\mathrm{a}) = (\gamma(\mathrm{C}, \mathrm{D}) - \gamma(\mathrm{C} - \{\mathrm{a}\}, \mathrm{D}))/\gamma(\mathrm{C}, \mathrm{D}) = 1 - (\gamma(\mathrm{C} - \{\mathrm{a}\}, \mathrm{D}) / \gamma(\mathrm{C}, \mathrm{D}))$$

And it is simply denoted by $\sigma(\mathrm{a})$, where $0 \leq \sigma(\mathrm{a}) \leq 1$.

2.3 Bayesian Decision Theoretic Rough Set [1, 2]

Let, D_{POS} denote the positive region in BDTRS model. For an equivalence class $[\mathrm{x}]_\mathrm{C} \in \pi_\mathrm{A}$

$$D_{POS}([x]_C) = \{ D_i \in \pi_D \colon P(D_i/[x]_C) > P(D_i)\}$$

For equivalence classes $[x]_C$ and $[y]_C$ the elements of a positive decision-based discernibility matrix, $M_{D_{POS}}$ is defined as follows.

$$M_{D_{POS}}([x]_C, [y]_C) = \{a \in C \colon I_a(x) \neq I_a(y) \wedge D_{POS}([x]_C \neq D_{POS}([y]_C)\}$$

Attribute reducts are the set of prime implicants of the reduced disjunctive form of the discernibility function. So a positive decision reduct is a prime implicant of the reduced disjunctive form of the discernibility function.

$$f(M_{D_{POS}}) = \wedge \{ \vee (M_{D_{POS}}([x]_C, [y]_C)) : \forall x, y \in U$$
$$(M_{D_{POS}}([x]_C, [y]_C) \neq \phi)\}$$

In order to derive the reduced disjunctive form, the discernibility function $f(M_{D_{POS}})$ is transformed by using the absorption and distributive laws. Accordingly, finding the set of reducts can be modeled based on the manipulation of a Boolean function.

3 The BDTRS Model

3.1 Processing Steps

Computation of Decision Table from raw dataset can be performed using basic R functions (e.g., read.table(), SF.asDecisionTable()) as shown in [9]. Figure 1, shows the actual steps of attribute reduction process. The first R function, BayesDtPos(), computes Bayesian Decision Theoretic Positive region. Algorithm (Algorithm-1) for this function is given in Sect. 3.2. Input to this function is a decision table. Next, Bayesian Decision Theoretic Discernibility Matrix is computed by the function named as BayesDtDisMat(). It accepts a Positive Region as shown in Fig. 1. Algorithm (Algorithm-2) for this function is also given in Sect. 3.2. Finally, FS. one.reduct.computation() [9] computes reduced attribute set.

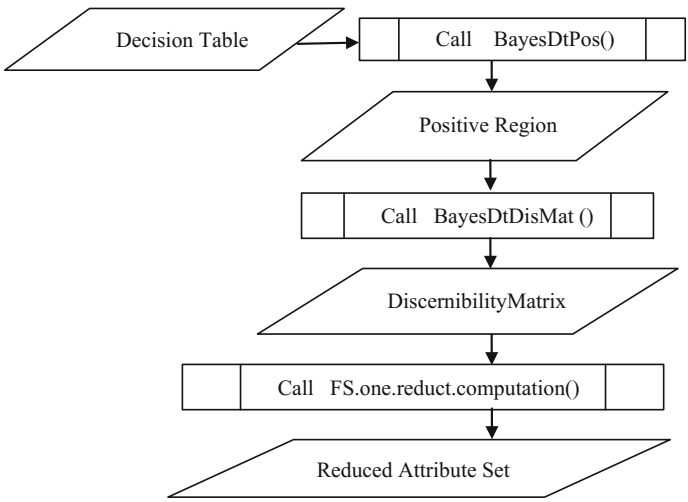


Fig. 1 Steps of attribute reduction using BDTRS model

3.2 Algorithms

Algorithm-1. **BayesDtPos (DT)**:

Given a Decision Table (DT), this algorithm computes Bayesian Decision Theoretic Positive Region (D_{pos}). Equivalent classes (Indiscernibility relation), computed from DT, are represented as [IND]. The vector D, represents the unique decisions (decision column) of the DT. i, j, X_1, X_2, X, n are used as temporary variables.

1. Read a Decision Table as DT.
2. Compute [IND], Indiscernible relation/Equivalent classes from DT.
3. Compute unique decision set from decision column of DT into D.
4. Repeat step 4 for [IND_i], i = i to n, n being the number of Equivalent classes.
5. Repeat steps 5.1 to 5.4 for each Decision (D_j), j = 1 to k, k being the number of unique decision.

   5.1. X_1 = Number of objects in [IND_i], having D_j as decision.
   5.2 X_2 = Number of objects in [IND_i].
   5.3 $X = X_1/X_2$.
   5.4 If ($X > P(D_j)$) [$P(D_j)$ is the probability of $D_{j\text{th}}$ Decision]

      5.4.1 Assign Decision D_j, to $D_{pos,}$ corresponding to each object of [IND_i].
6. Return (D_{pos}).

Algorithm-2. **BayesDtDisMat (DT)**:

The variables DT and D_{pos} are already described in Algorithm-1. dataMatrix is a $n \times m$ matrix, where n represents number of objects in DT and m represents number of condition attributes in DT. decVector is a $n \times 1$ vector. objIdx, objIdx1, objIdx2 are temporary variables. discernibilityMatrix is a $n \times n$ matrix, where n described as above. attribute list is a character vector for storing attribute names of the condition attributes. This algorithm returns Discernibility Matrix.

1. Call Algorithm-1 with DT as an argument to compute D_{pos}.
   [i.e. D_{pos}= BayesDtPos (DT).]
2. dataMatrix = DT – {Decision column of DT}
3. decVector = Decision column of DT.
4. Repeat for ith object in DT [i.e., i = 1 to (n − 1), n is number of objects in DT]

   4.1. objIdx1 = (decVector[i] != (decVector[i + 1] to decVector[n]))
   4.2. objIdx2 = (D_{pos}[i] != (D_{pos}[i + 1] to D_{pos}[n]))
   4.3. objIdx = {objIdx1 ∩ objIdx2}
   4.4. Repeat for jth object in objIdx [i.e., j = 1 to number of objects in objIdx]

      4.4.1 attributeList = Column names of dataMatrix where Column name of ith object != Column name of jth object.
      4.4.2 discernibilityMatrix[j, i] = attributeList.
      [i.e., Assign the attributeList to discernibilityMatrix]

5. Return (discernibilityMatrix)

3.3 Execution of BDTRS Based Functions in R Environment

R commands to execute our BayesDtPos() and BayesDtDisMat() functions using *housing* [4] dataset (https://archive.ics.uci.edu/ml/datasets/housing), is shown in this section (R version 3.2.2 has been used). Installation procedure of R and relevant packages ("RoughSets", "stringi" etc.) is available in CRAN at http://cran.r-project.org.

```
>HousingFrame = read.table ("housing.csv", header = TRUE, sep = ",")
>HousingDecisionTable = SF.asDecisionTable (HousingFrame, decision.attr = 14, indx.nominal = c(1:13))
>BDTRS_PossitiveRegion = BayesDtPos (HousingDecisionTable, c(1:13))
>BDTRS_DisMatrix=BayesDtDisMat(HousingDecisionTable, BDTRS_PossitiveRegion, range.object = NULL, return. matrix =TRUE)
>BDTRS_AttributeSubSet = FS.one.reduct.computation (BDTRS_DisMatrix)
>BDTRS_AttributeSubSet
$decision.reduct
$decision.reduct[[1]]
[1] "RAD" "RM" "DIS" "B" "AGE" "CRIM" "CHAS" "LSTAT" "TAX"
```

4 Results of Comparative Study

The BDTRS model (implemented as BayesDtPos() and BayesDtDisMat()) for attribute reduction is compared with two popular rough set based R functions: FS.quickreduct.RST() [9, 10] and FS.greedy.heuristic.reduct.RST() [9, 11, 13] (implemented in "RoughSets" package of R). For this, benchmark datasets such as *wine* [3] (available at https://archive.ics.uci.edu/ml/datasets/wine) and *housing* [4] (available at https://archive.ics.uci.edu/ml/datasets/housing) are used. The functions and computed output on *wine* dataset and *housing* dataset is shown in Table 1 and Table 2 respectively. Attribute significance, computed in R environment is based on the method described in Sect. 2.2.

Observations from Table 1: Attributes with low significance (significance <0.5) such as "malid_acid" (significance: 0.24) and "od" (significance: 0.21) are included by FS.quickreduct.RST() and FS.greedy.heuristic.reduct.RST() respectively. BDTRS Model does not include such insignificant attributes

Observations from Table 2: BDTRS Model includes only 9 (nine) significant attributes (Significance >0.5) where as the other two methods includes 10 (ten) attributes each. Moreover, the function FS.quickreduct.RST() and FS.greedy.heuristic.reduct.RST() includes insignificant attributes (Significance <0.5), "PTRATIO" and "INDUS" respectively.

Table 1 Results computed by various rough set based R functions on *Wine* dataset

Attribute reduction methods	Computed reduced attribute set
BDTRS Model	{color_intensity, proline, flavanoids, hue, alcohol}
FS.quickreduct.RST()	{alcohol, malid_acid, flavanoids, color_intensity, proline}
FS.greedy.heuristic.reduct.RST ()	{alcohol, flavanoids, hue, od, proline}
Computed tribute significance in ascending order: "proline" (significance: 0.7), "color_intensity" (significance: 0.67), "flavanoids" (significance: 0.62), "hue" (significance: 0.61), "alcohol" (significance: 0.54)	

Table 2 Results computed by various rough set based R functions on *housing* dataset

Attribute reduction methods	Computed reduced attribute set
BDTRS model	{RAD, RM, DIS, B, AGE, CRIM, CHAS, LSTAT, TAX}
FS.quickreduct.RST()	{RAD, RM, DIS, B, AGE, CRIM, CHAS, PTRATIO, NOX, TAX}
FS.greedy.heuristic.reduct. RST()	{RAD, RM, DIS, B, AGE, CRIM, CHAS, LSTAT, NOX, INDUS}
Computed tribute significance in ascending order: B (significance: 0.75), RAD (significance: 0.74), RM (significance: 0.74), LSTAT (significance: 0.70), AGE (significance: 0.68), CHAS (significance: 0.65), CRIM (significance: 0.63), TAX (significance: 0.60), DIS (significance: 0.55), ZN (significance: 0.44), NOX (significance: 0.32), PTRATIO (significance: 0.22), INDUS (0.14).	

5 Conclusion

In this paper, various implementation aspects such as algorithms, processing steps, commands to execute the BDTRS model (functions) in R environment are presented using example. The implemented model is also compared with few existing rough set based attribute reduction functions. Experimental results show that BDTRS model gives improved result over the other two attribute reduction methods as it includes all the significant attributes (significance >0.5), leaving out all the insignificant attributes (significance <0.5), in the reduced attribute set.

In future, the BDTRS model may be used to deal with specific data mining problems, such as BDTRS based Decision Tree induction for classification, constructing BDTRS based Bayesian classifier, etc.

Acknowledgements: This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors. Utpal Pal would like to express his gratitude to the Department of Mathematics and Department of IT, Tripura University (A Central University), Tripura, India for supporting him to pursue the Ph.D. program.

References

1. Halder, S.B.: A Study on bayesian decision theoretic rough set. Int. J. Rough Sets Data Anal. (IJRSDA) 1.1, 1–14 (2014)
2. Halder, S.B., Debnath, K.: Attribute reduction using bayesian decision theoretic rough set models. Int. J. Rough Sets Data Anal. (IJRSDA) 1.1, 15–31 (2014)
3. Forina, M., Lear, R., Armanino, C., Lauter, S.: PARVUS—an extendible package for data exploration, classification and correlation. J. Chemonetrics **4**(2), 191–193 (1988)
4. Harrison, D., Rubinfeld, D.L.: Hedonic prices and the demand for clean air. J. Environ. Econ. Manage. **5**, 81–102 (1978)
5. Ihaka, R., Gentleman, R.R.: A language for data analysis and graphics. J. Comput. Graph. Stat. **5**, 299–314 (1996)
6. Li, C., Yang, Y., Jia, M., Zhang, Y., Yu, X., Wang, C.: Phylogenetic analysis of DNA sequences based on k-word and rough set theory. Phys. A Stat. Mech Appl. **398,** 162–171 (2014)
7. Pawlak, Z.: Rough sets. Int. J. Comput. Inform. Sci. **11**, 341–356 (1982)
8. R Development Core Team, R: A Language and Environment for Statistical Computing. The R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/ (2011). Accessed 08 June 2016. ISBN: 3-900051-07-0
9. Riza, L.S., Janusz, A., Bergmeir, C., Cornelis, C., Herrera, F., Slezak, D., Benitez, J.M.: Implementing algorithms of rough set theory and fuzzy rough set theory in the R package ''RoughSets''. Inf. Sci. (ELSEVIER) **287**, 68–89 (2014)
10. Shen, Q., Chouchoulas, A.: A modular approach to generating fuzzy rules with reduced attributes for the monitoring of complex systems. Eng. Appl. Artif. Intell. **13**, 263–278 (2000)
11. Slezak, D.: Approximate Entropy Reducts. Fundam. Informaticae **53**(3–4), 365–390 (2002)
12. Slezak, D., Ziarko, W.: Bayesian rough set model. In: Proceedings of the International Workshop on Foundation of Data mining, Japan, pp. 131–135. 9 Dec 2002
13. Wroblewski, J.: Ensembles of classifiers based on approximate reducts. Fundam. Informaticae **47**(3–4), 351–360 (2001)
14. Yao, Y.Y., Wong, S.K., Lingras, P.: A decision theoretic rough set model. In: Ras, Z.W., Zemankova, M., Emrich. M.L. (eds.) Methodologies for intelligent systems, vol. 5, pp. 17–24. North Holland, New York, (1990)
15. Ziarko, W.: Variable precision rough set model. J. Comput. Syst. Sci. **46**, 39–59 (1993)

CBSTD: A Cloud Based Symbol Table Driven DNA Compression Algorithm

Annwesha Banerjee Majumder and Somsubhra Gupta

Abstract In this paper, we propose symbol table driven DNA compression algorithm aimed to use as a cloud service. Bioinformatics requires a huge amount of genomic data to analysis, so optimal storage and compression is a great challenge to this field. We categorized the DNA sequence into three different parts according to the occurrence of A T C and G and use two different symbols tables to map the DNA sequences into a compressed sequence. We are intended to deploy our proposed compression algorithm in cloud, so that the user of this field can access this Software as a Service over the cloud. Through our proposed method of compression, we claim to achieve a compression rate of 1.82.

Keywords Bioinformatics · Compression algorithm · DNA sequence · Compression rate · Cloud

1 Introduction

Bioinformatics basically is the field of informatics and biology. It is the interaction between computation and biology where computation is being used to biological data analysis and at the same time machine learning is one of the basic requirement for biological data computation [1]. Biological computation is database oriented, there are several public biological databases available such as GenBank, DDBJ, and EMBL for nucleic acid sequences [2].

Around 3 billion characters and over 23 pair of chromosome are there in human genome so the database that stores the genome information are very huge.

A.B. Majumder (✉) · S. Gupta
JIS College of Engineering, Kalyani, West Bengal, India
e-mail: annwesha.banerjee@gmail.com

S. Gupta
e-mail: gsomsubhra@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_45

DNA: DNA is treated as "blue print of life". It contains all the information to create life. DNA contains the information needed to create the amino acids sequences of proteins. The unit of building block of DNA Adenine (A), Cytosine (C), Guanine (G), and Thymine (T) are the four bases in DNA. A pair with T that is 2H bond and C pairs with G that is 3H bond. The double helix of DNA controls heredity on molecular level. The heredity information is the stored sequence of bases with a message in form of only four letters A, T, C, and G. DNA preserves and uses this information [3]. DNA is the storage repository for how cell function, [4].

DNA sequences have following characteristics:

1. Contain off-repeated substring
2. Some strings are palindrome
3. Some strings are reverse palindrome [5].

Gene: A gene is a specific sequence of nucleotide bases, whose sequences carry the information required for constructing proteins, which provide the structural components of cells and tissues as well as enzymes for essential biochemical reactions. The human genome is estimated to comprise more than 30,000 genes.

Protein: Protein is a linear sequence of amino acids, shown in the Fig. 1 form a very long chain via peptide linkage. Gene is a segment of DNA which contains the genetic code to make a certain protein or part of protein.

Inheritance pattern is the predictable pattern seen in the transmission of genes from one generation to the next and their expression in the organism that possesses them.

Bioinformatics works basically with large data set DNA sequence, researchers focus in the field of compression. Compression of DNA sequences can be achieved by the help of efficient pattern recognition strategies. Compression basically reduces the space when storing them and reduces time when transmitting them. Compression algorithm must follow the following:

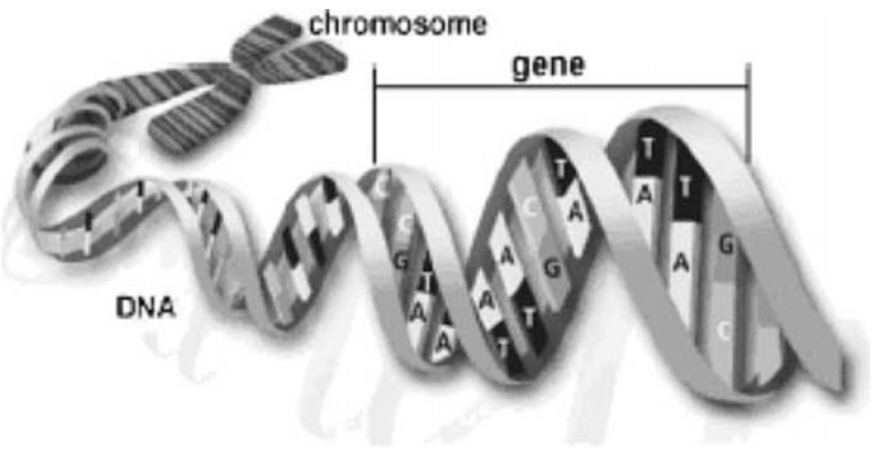


Fig. 1 DNA double helix and gene sequence

$$(\#\text{ bytes after compression}/\#\text{ bytes before compression}) \leq \alpha \quad (1)$$

with $\alpha > 0$ as small as possible.

Redundancy is very common in DNA sequences, and compression needs to avoid redundancies. Different compression algorithms are already been introduced and here as we proposed text compression of DNA sequence; so we mainly discussed about DNA compression algorithms.

Several DNA sequence compression algorithm have already been proposed and implemented. BioCompress -1 and BioCompress -2 [6, 7] proposed by Grumbach and Tahi used Ziv-and-Lempel data compression method. Theses methods can find the replica of sequences, such as repeats, palindromes, and complementary palindromes. CFACT E.RIVAL 1996 [8] detects repeats using suffix tree data structure. It searches longest exact matching repeats. Another interesting algorithm GenCompress [7] was proposed by Chen et al. which can handle approximate and inexact repeats. Matsumoto et al. proposed Context-tree weighting [9, 10] that is capable of handling both short and long repeats. GenomeCompress [11–13] is the algorithm that divides the total sequence into four bits long segments and encode these into five bits long segment and it compresses both repeat and non-repeat sequences. Chen et al. have also proposed an algorithm that finds all approximate repeats by help of Pattern Hunter, named DNA Compress [6, 7, 14]. HUFFBIT COMPRESS [15] is based on extended binary tree for compression and its compression rate is 1.006 bits per phase and HASHBASED [16] initially builds a hash table and assign an unique character to each of the factor that works as hash key and these keys are assigned to four bit long subparts of the sequences. LOOK UP TABLE [17–19] is based on fixed length LUT and LZ77. Differential Direct Coding [20] this algorithm accommodates large data sets, consist of multiple sequences and auxiliary data. LSDB METHOD [21] performs gene-based compression and reduce memory requirement. Panneer Arokiaraj S et al. have proposed an algorithm An Improvised DNA Sequence Compressor to identify different pattern, repeating patterns are encoded based on their uniqueness and other patterns are identified and stored in symbol table [22]. M. Mary, Shanthi Rani has proposed a reference based DNA Compression Algorithm [23].

Bioinformatics complex computation can be achieved through distributed computing. Genomic research requires the dynamic computation of large dataset. Cloud computing which is an on demand, pay per use, IT service centric dynamic approach is now trying to provide some assistance to the bioinformatics research domain. Hadoop is now being used to store huge genomic data and perform data analysis. Cloud computing is so popular because of its cost effectiveness [24]. Clouds that provide bioinformatics services are data dependent where annual worldwide sequencing capacity is beyond 13 Pbp. Due to these requirement Data As a Service is utmost important [25, 26] (Fig. 2).

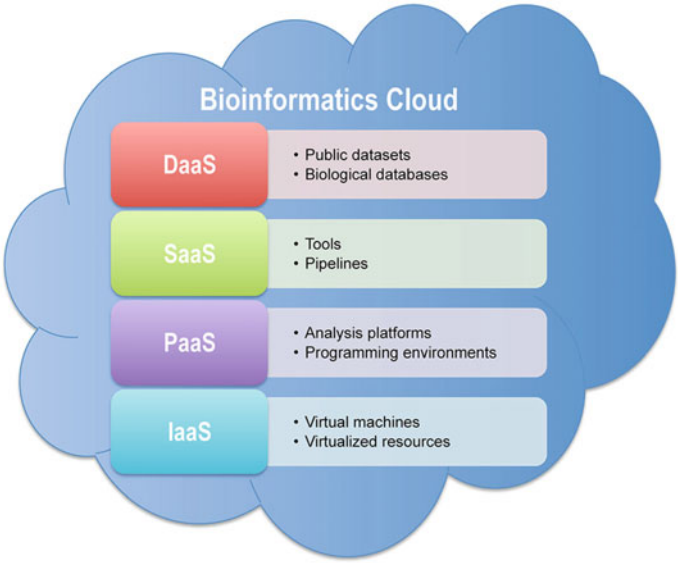


Fig. 2 Bioinformatics cloud

Amazon Web Services (AWS) provides a centralized repository of public data sets, GenBank, Ensembl, 1000 Genomes, Model Organism Encyclopedia of DNA Elements, Unigene, Influenza Virus, etc. AWS contains multiple public datasets for a variety of scientific fields, such as biology, astronomy, chemistry, climate, economics, etc. All public datasets in AWSare delivered as services and thus can be seamlessly integrated into cloud-based applications [27].

2 Proposed Method

To compress DNA sequence, we use substation table for mapping different pattern to different symbols. We proposed to store the DNA sequence in compressed form in cloud and users can access the Software as a Service over the cloud with the advantages of cloud computing dynamic on demand pay per use services and at the same time the user machine would not be congested by huge data and complex computation and multiple user can access the application in scalable manner. The compression rate we achieved through the proposed method is on average of 1.82.

We can brief our proposed method as the following:

Step 1: Compression
Step 2: Decompression
Step 3: Deployment of Application over the cloud

2.1 Compression

In our proposed method, we use substitution of DNA set D = {d_1, d_2, d_3, …, d_n} to set of compressed symbols S = {S1, S2}, where di = 1, 2, …, n = {A, T, C, G} and S1 = {a, b, c, … z,} and S2 = {!, @, #, …, +}

$$f(\mathrm{D} \rightarrow \mathrm{S}) \tag{2}$$

Basically our algorithm subdivides the total DNA sequence into four-length block and then look for following three pattern of string:

a. All four symbols are different.
   D1 = {d_1, d_2, d_3, d_4: all the ds are distinct}
b. Two nonconsecutive symbols are same
   D2 = {d_1, d_2, d_3, d_4: $d_1 = d_3$ or $d_1 = d_4$ or $d_2 = d_4$}
c. Consecutive symbols are same
   $D3 = D \cap (D1 \cup D2)$
   D3 = {d_1, d_2, d_3, d_4: any two or more consecutives symbols are same}

A. Creation of Symbol Table

a. Take four bits together, then create symbol table with 24, **S1 = {a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z}** combinations where no symbols are repeating and assign a to x as encoded symbols,
b. Take two (non-repeating) symbols together and assign symbols from S2 = {t, \t, b, \b, n, \n, r, \r, f, \f, +, −} for all 12 combinations.

B. Process of Compression

1. Divide the sequence into sub-parts of length 4
   a. If all the symbols are different as the symbols belong to **D1 = {d_1, d_2, d_3, d_4: all the ds are distinct}** then,

$$f(\mathrm{D1} \rightarrow \mathrm{S1}) \tag{3}$$

   b. If two nonconsecutive symbols are same as they belong to set **D2 = {d_1, d_2, d_3, d_4: $d_1 = d_3$ or $d_1 = d_4$ or $d_2 = d_4$}** then divide the sub sequence into parts having two symbols and encode with

$$f(\mathrm{D2} \rightarrow \mathrm{S2}) \tag{4}$$

   c. If consecutive symbols are same simple replace by symbols and no of occurrence.
   D3 = {d_1, d_2, d_3, d_4: any two or more consecutives symbols are same} and let S3 be the set to which the D3 will be encoded **S3 = {$d_1$4, $d_1$3d_4, $d_1$2$d_3$$d_4$, $d_1$$d_2$2$d_4$, $d_1$$d_2$3, $d_1$$d_2$2$d_4$, $d_1$$d_2$$d_3$2, $d_1$$d_4$3, $d_1$2$d_3$2}**

Then mapping will be done by following:

$$f(\mathrm{D3} \rightarrow \mathrm{S3}) \tag{5}$$

2.2 Decompression

Through the decompression, the actual DNA sequence is reverted from the encoded sequence.

1. Read the sequence.
2. If the read symbol belonging to set S1 = {a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z} then decode it to set **D1 = {d_1, d_2, d_3, d_4: all the ds are distinct}** consulting the symbol Table 1

$$f'(\mathrm{S1} \rightarrow \mathrm{D1}) \tag{6}$$

Table 1 Symbol Table 1

DNA sequence	Encoded symbols
ATCG	a
ATGC	b
AGTC	c
AGCT	d
ACGT	e
ACTG	f
CATG	g
CAGT	h
CTAG	i
CTGA	j
CGTA	k
CGAT	l
TCAG	m
TCGA	n
TACG	o
TAGC	p
TGCA	q
TGAC	r
GATC	s
GACT	t
GCAT	u
GCTA	v
GTAC	w
GTCA	x

Table 2 Symbol Table 2

DNA sequence	Encoded symbols
AT	t
TA	\t
AG	b
GA	\b
AC	n
CA	\n
CT	r
TC	\r
GC	f
CG	\f
GT	+
TG	–

3. If the read symbols belonging to set S2 = {t, \t, b, \b, n, \n, r, \r, f, \f, +, −} then the symbols are to be mapped to set **D2 = {d_1, d_2, d_3, d_4: $d_1 = d_3$ or $d_1 = d_4$ or $d_2 = d_4$}** by consulting the symbol Table 2.

$$f'(\mathrm{S2} \rightarrow \mathrm{D2}) \tag{7}$$

4. If the read symbols are belonging to set **S3 = {$d_1$4, $d_1$3d_4, $d_1$2d_3d_4, $d_1d_2$2d_4, $d_1d_2$3, $d_1d_2$2d_4, $d_1d_2d_3$2, $d_1d_4$3, $d_1$2$d_3$2}**, then map these to set
5. **D3 = {d_1, d_2, d_3, d_4**: any two or more consecutives symbols are same}

$$f'(\mathrm{S3} \rightarrow \mathrm{D3}) \tag{8}$$

2.3 *Deploying the Application in Cloud*

The DNA sequence are stored in a compressed manner in the cloud that minimizes the storage requirements, and the client can access the sequence over the Internet. We have implemented the application using Java Web application framework where the compression and decompression is implemented in Plain Old Java Class which is reusable and platform independent are able to execute anywhere (Figs. 3 and 4).

Following table describes a comparative analysis of our proposed algorithm with GeneCompression and DNA Compression algorithm [28].

Comparison with Existing Algorithm

See Table 3.

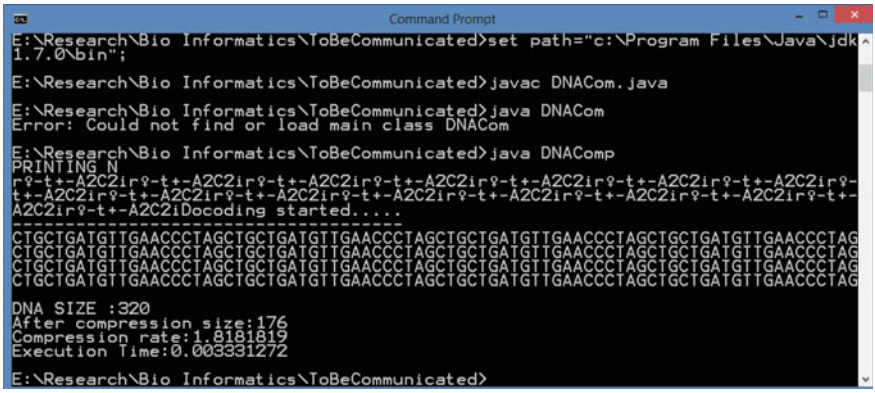


Fig. 3 Test Result1

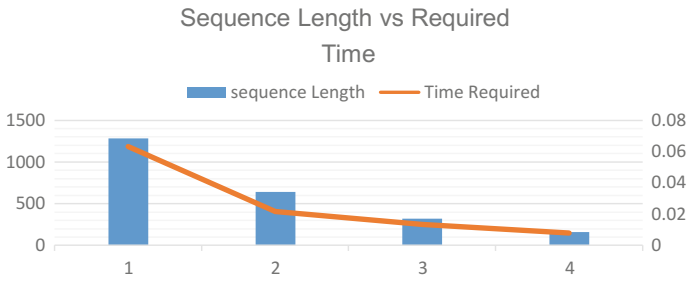


Fig. 4 Variation of time for different sequence length

Table 3 Comparison with some existing algorithm

Compression algorithm	Compression ratio
GeneCompression	1.7428
DNACompression	1.7254
Proposed CBSTD	1.82

3 Conclusion

In this paper, we have proposed a method to compress DNA sequence which would be provided as DaaS (Data as a service) over cloud which can be easily accessible from anywhere any time in the computation of biological data. We have achieved compression rate of 1.82 and that can be improved by increasing the block size of symbol mappings.

References

1. Baldi, P., Brunak, S.: Bioinformatics: The Machine Learning Approach. MIT Press, Cambridge, MA (1998)
2. Barker, W.C., Garavelli, J.S., Huang, H., McGarvey, P.B., Orcutt, B., Srinivasarao, G.Y., Xiao, C., Yeh, L.S., Ledley, R.S., Janda, J.F., Pfeiffer, F., Mewes, H.W., Tsugita, A., Wu, C.: The protein information resource (PIR). Nucleic Acids Res. **28**, 41–44 (2000); Bairoch, A., Apweiler, R.: The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res. **28**, 45–48 (2000)
3. Wong, L.: Some new results and tools for protein function prediction, RNA target site prediction, genotype calling, environmental genomics, and more. J. Bioinform. Comput. Biol. **9**(6) (2011)
4. Lewin, B.: Genes VII. Oxford University Press, New York, NY (1995)
5. Manzini, G., Rastero, M.: A simple and fast DNA compressor, software: practice and experience. MIUR Support Projects (ALINWEB) **34**(14), 1397–1411 (2004)
6. Chen, X., et al.: DNA compress: fast and effective DNA sequence compression. Bioinform. Appli. Note **18**(12), 1696–1698 (2002)
7. Textual data compression in computational biology: a synopsis Raffaele Giancarlo∗. Davide Scaturro Filippo Utro **25**(13), 1575–1586 (2009). doi:10.1093/bioinformatics/btp117
8. Rivals, E., Delahaye, J.-P., Dauchet, M., Delgrange, O.: A guaranteed compression scheme for repetitive DNA sequences. LIFL Lille I University, Technical report IT-285 (1995)
9. Bao, S., et al.: A DNA Sequence Compression Algorithm Based on LUT and LZ77 (2005)
10. Ziv, J., Lempel, A.: A universal algorithm for sequential data compression. IEEE Trans. Inform. Theory **IT-23** (1977)
11. Ghoshdastider, U., et al.: GenomeCompress: A Novel Algorithm for DNA Compression (2005). ISSN 0973–6824
12. Chen, Xin, Kwong, Sam, Li, Ming: A compression algorithm for DNA sequences and its applications in genome comparison. Genome Inform. **10**, 51–61 (1999)
13. Matsumoto, T., et al.: Biological sequence compression algorithms. Genome Inform. **11**, 43–52 (2000)
14. Chen, X., Li, M., Ma, B., Tromp, J.: DNACompress: fast and effective dna sequence compression. Bioinformatics **18** (2002)
15. Raja Rajeswari, P., Apparao, A., Kiran Kumar, R.: HUFFBIT COMPRESS—algorithm to compress DNA sequences using extended binary trees. J. Theor. Appl. Inform. Technol. 101–106 (2005–2010)
16. Mehta, A., et al.: DNA compression using hash based data structure. IJIT&KM **2**(2), 383–386 (2010)
17. Bharti, R.K., et al.: A biological sequence compression based on approximate repeat using variable length LUT. Int. J. Adv. Sci. Technol. **3**(3), 71–75 (2011)
18. Bharti, R.K., et al.: Biological sequence compression based on cross chromosomal properties using variable length LUT. CSC J. **4**(6), 217–223 (2011)
19. Bharti, R.K., et al.: Biological sequence compression based on properties unique and repeated repeats using variable length LUT. CiiT J. **3**(4), 158–162 (2011)
20. Bolshoy, A.: DNA sequence analysis linguistic tools: contrast vocabularies, compositional spectra and linguistic complexity. Appl. Bioinform. **2**, 103–112 (2003)
21. Wu, C.P.P., et al.: Cross chromosomal similarity for DNA sequence compression. Bioinformatics **2**(9), 412–416 (2008)
22. Panneer Arokiaraj, S., Robert, L.: An improvised DNA sequence compressor using pattern recognition. Int. J. Eng. Technol. (IJET) **5**(6), (Dec 2013–Jan 2014). ISSN 0975-4024
23. Mary Shanthi Rani, M.: A new referential method for compressing genomes. Int. J. Comput. Bioinform. Silico Model. **4**(1), 592–596 (2015)
24. Introduction to Cloud Computing Fact Sheet

25. Truong, H.L., Dustdar, S.: On analyzing and specifying concerns for data as a service. In: 2009 IEEE Asia-Pacific Services Computing Conference (APSCC2009), pp. 83–90 (2009)
26. DaaS: The New Information Goldmine. http://online.wsj.com/article/SB125071202052143965.html
27. Fusaro, V.A., Patil, P., Gafni, E., Wall, D.P., Tonellato, P.J.: Biomedical cloud computing with amazon web services. PLoS Comput. Biol. **7**(8), e1002147 (2011)
28. Rajarajeswari, P., Apparao, A.: DNABIT compress—genome compression algorithm. Bioinformation **5**(8), 350–360 (2011) (Published online 22 Jan 2011)

Optimization of E-Jet Based Micro-manufacturing Process Using Desirability Function Analysis

Raju Das, Amit Kumar Ball and Shibendu Shekhar Roy

Abstract Electrohydrodynamic (EHD) printing is a micro- and nano-manufacturing process of printing high-resolution functional material on a substrate. It is a very exciting alternative to the conventional inkjet printing technology for micro-droplet generation. In this work, an approach has been made to tune the process control parameters to achieve better functioning of the printing process. The droplet size and the printing frequency have been taken as performance measure of the printing process whereas applied voltage, back pressure, and nozzle standoff height have been selected as the process parameters which are to be tuned through optimization. Desirability function analysis have been employed to optimize the process parameters for multiple output variables simultaneously. Composite desirability values have been computed and based on these values; the optimal process parameters which leads to better printing performance have been proposed.

Keywords Micro-manufacturing · Electrohydrodynamic · Inkjet printing · Desirability function analysis (DFA)

1 Introduction

With growing demand for micro- and nano-level devices to meet the consumer needs, search for alternative technologies to achieve micro- and even nano-scale fabrication has also been intensified. Several researchers have devoted their time to explore new technologies to meet today's challenges in design and manufacturing arena to crater the aspiration of end users. As with any subtractive manufacturing process, excess material has to be removed to get the desired product. So material waste is a pertinent issue with this kind of fabrication process. But in case of additive manufacturing scenario, virtually no material has been wasted during fabrication. In

R. Das · A.K. Ball · S.S. Roy (✉)
Department of Mechanical Engineering, National Institute of Technology, Durgapur, Durgapur 713209, India
e-mail: ssroy99@yahoo.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_46

many applications delivering the expensive functional material is upmost necessity. In those applications, if the subtractive manufacturing approach would be adopted, most of the resources will be wasted. Electrohydrodynamic printing is a versatile micro- and nano-manufacturing technology for micro-dispensing of functional material. It quietly becomes a suitable candidate as an alternative of lithography technique because of its inherent advantages such as it is a noncontact mask less additive in nature micro/nano-fabrication technology where no material is wasted. It finds its application in vast areas like flexible printed electronics, biosensors, microelectromechanical systems (MEMS), photovoltaic, drug development etc. [1]. It also provides greater resolution feature than its conventional inkjet technology peers. The minimum resolution that can be achieved with conventional inkjet technology is limited to 20–30 μm [2]. The above-mentioned limitation can greatly overcome by EHD printing.

In EHD printing, an electric field is applied between the conducting nozzle (ejection aperture) and the substrate (the surface on which printing is done). Due to the applied electric field the liquid meniscus which is accumulated at the tip of the nozzle deformed into conical shape which is known as Taylor's cone [3]. When the applied filed reaches it is critical value, i.e., sufficient enough to overcome the surface tension force which is acting at the tip of the nozzle, ink material eventually coming out of the nozzle. The printing operation depends on the voltage difference between the nozzle and the substrate, the standoff height, nozzle diameter, applied pressure to the ink chamber, viscosity of the ink type and other factors. In this technique the liquid ink material gets ejected at the tip of the Taylor cone rather than the tip of the nozzle. This is the main reason behind this technology's ability to print droplets whose size is very much smaller than the nozzle diameter. A very high resolution of printing of different functional material can be achieved by this technology [4]. Figure 1 shows the schematic diagram of the EHD printing system.

As can be seen from Fig. 1 the main elements for E-jet printing include an ink chamber, nozzle tip, substrate and positioning system (stage). A whole operation consist of four stages namely stages of liquid accumulation, cone formation, droplet ejection, and relaxation.

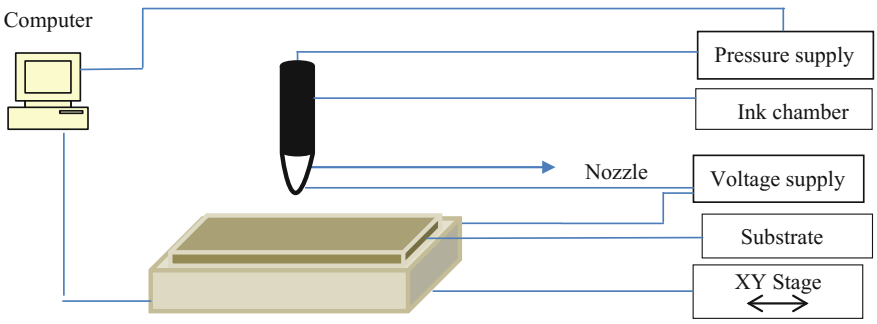


Fig. 1 Schematic diagram of the EHD printing system

The disintegration of micro-droplet due to applied electrical field was discovered long ago [5]. After that several researchers have studied that principle to print high-resolution functional material droplets. Chen et al. [6] printed particles which are much smaller than the nozzle diameter using EHD DOD technique, thus eliminating the difficulty to manufacture micro-size nozzle and proves its employability. Many researchers have shown the capabilities of E-jet printing applied to different applications to print different kind of functional material [7–10]. Park et al. [4] describes high-resolution electrohydrodynamically printed pattern and functional devices with submicron resolution. They have used direct high-speed imaging to capture droplet formation processes. They have produced droplets with both pulsating jet mode and stable jet mode. Chen et al. [11] and Choi et al. [12] have proposed scaling laws of different parameters for EHD printing. They showed that the jet diameter is proportional to the square root of the nozzle size and inversely proportional to the electric field strength. Printing speed in EHD printing is comparatively low to that of the conventional printing methods. Kim et al. [13] proposed a hybrid printing methodology where, they had employed piezoelectric actuator to supply the ink at the nozzle tip and then applied electric field for printing. Mishra et al. [14] demonstrated high-speed high accuracy drop on demand printing mode using pulsed DC voltage. Barton et al. [15] of University of Illinois developed a compact E-jet printing module aiming to provide compact, affordable, and user-friendly E-jet printing system. Many researchers also take up the issue of heterogeneous printing of functional material with the help of multi-nozzle print head [16].

In the literature cited above, no work has been attempted to optimize the process parameters of EHD printing system. Further, to best of our knowledge, any statistical analysis has also not been carried out yet. This paper presents Desirability functional analysis (DFA) approach to find out optimal setting of the control variables so as to attain better printing performance.

The context is organized in the following manner: brief description about methodology followed by results and discussion about the conducted analysis presented. Finally, the paper is concluded with the findings of this study.

2 Proposed Methodology

The proposed methodology utilizes Desirability function analysis (DFA) to improve the printing function of functional material. It is a quite simple but capable enough of handling multi criteria objectives problems [17]. A brief introduction of above-mentioned methodology is given in the subsequent sections.

2.1 Desirability Functional Analysis (DFA)

DFA is a very simple and widely used multi-objective optimization technique which involves conversion of multi-response characteristics into a single response characteristic called composite desirability index [18]. Thus the multi-objective optimization problem can easily be converted into single objective problem. This method provides great help to identify the operating conditions that leads to 'most desirable' responses. Optimization of composite desirability index leads to simultaneous optimization of both the response parameters in this study namely droplet diameter and printing frequency. In this approach first the individual response parameters are converted into a range of 0 and 1 with the help of a transfer function known as desirability function 'd_i' ($0 < d_i < 1$). Higher d_i values indicates response is more desirable and may approach to the ideal condition, i.e., 'd_i' = 1, and if 'd_i' = 0; this implies that response is undesirable and lies outside the acceptable limit. After computing the desirability index for response variables, composite desirability index 'D' have to be computed to make the decision. Higher the value of 'D' indicates input parameters setting is in optimal condition [19]. The following steps have been followed while conducting DFA:

1. Choose an appropriate transfer function according to the goal
2. Calculate individual desirability index for response variable 'd_i'
3. Compute composite desirability index 'D'
4. Identify process parameter settings corresponding to maximum 'D' value.

2.2 Desirability Function

There are three desirability transfer function available namely 'larger the better' where goal is to maximize the response, 'smaller the best' where objective is to attain response value as small as possible and the third one is 'nominal is best'. This is used when attaining a prerequisite value is the target. In this study, 'smaller the best' have been chosen for droplet diameter and 'larger the better' has been employed for calculating desirability values of printing frequency. These transfer functions map the response variables into a unit less desirability value between 0 and 1. The following equations have been used to calculate 'd_i' values for the response variables.

For larger the better, the expression used for calculating the individual desirability index is given [20]

$$d_i = \begin{cases} 0, & x_i \le x_{\min} \\ \left(\frac{x_i - x_{\min}}{T - x_{\min}}\right)^r & y_{\min} \le x_i \le T, \quad r \ge 0 \\ 1, & x_i \ge x_{\min} \end{cases} \tag{1}$$

where 'x_i' represents response variable value at *i*th experimental run, T is the target value, in this study max value of x taken as the target value, and r is the weight. When 'x_i' exceeds a particular value, which can be viewed as a criterion, the desirability value becomes 1. And when 'x_i' falls below a particular value, which is unacceptable, the desirability index becomes 0.

For Smaller the best case scenario the following formula has been used:

$$d_i = \begin{cases} 1, & x_i \le x_{\min} \\ \left(\frac{x_i - x_{\max}}{T - x_{\max}}\right)^r & x_{\min} \le x_i \le T, \quad r \ge 0 \\ 0, & x_i \ge x_{\max} \end{cases} \tag{2}$$

When 'x_i' becomes min. of x the condition is desirable, i.e., d_i becomes 1 and when 'x_i' becomes greater than the max. of x it implies unfavorable condition and d_i becomes 0.

2.3 Composite Desirability Function

After calculating the individual desirability index values 'd_i' for all the responses, the next step is to calculate the composite desirability value 'D'. Higher the 'D' value represents more will be the desirability. The following formula have been employed to calculate D:

$$D = \sqrt[w]{d_1^{w_1} * d_2^{w_2} \ldots d_i^{w_i}} \tag{3}$$

where 'd_i' represents individual desirability and w_i is corresponding weight of the response variable. After getting the weight values the composite desirability index values have been calculated for each experimental run.

3 Results and Discussions

Experimental data with regard to different process parameters and printing performance are required for desirability function analysis (DFA). In this work, the experimental data have been taken from the previous studies conducted by the Graf [21]. The control parameters ware taken as-nozzle size, applied pressure, standoff height, and applied voltage. Two values of each of these variables were taken to conduct the experiments. The setting of standoff height includes 30–50 μm; those of applied back pressure include 0.25–1 psi; and the applied voltage is set of two values high and low. The particle analysis gave the average droplet diameter. Image J software has been employed for particle analysis. For calculating average droplet frequency, stage speed for each run and the travel distance of droplets has been used. In this study, we have taken applied pressure, standoff height, and

Table 1 Control variables and response variables [21]

Experiment no	Pres. (psi)	Gap (μm)	Volt. (V)	Ave. drop. freq. (Hz)	Ave. drop. dia. (μm)
1	1	30	245	50.88	1.63
2	1	30	260	383.6	2.07
3	1	50	275	74.08	1.32
4	1	50	300	326.6	2.16
5	0.25	30	267	83.38	1.79
6	0.25	30	285	285.1	1.47
7	0.25	50	297	39.77	2.19
8	0.25	50	315	253.1	1.33

applied voltage as control parameters and droplet diameter and droplet frequency as response variable. The desired droplet diameter should be as small as possible therefore the goal is minimum the better. As for droplet frequency desired target is larger the best. The observed values of various control parameters along with response variables are shown in Table 1. As discussed earlier in order to conduct desirability function analysis, first the individual response characteristics have to be transformed into a dimensionless utility function with the help of transfer function also known as desirability function. The individual desirability value 'd_i' for each response variable is calculated by Eqs. (1) and (2). Equation (1) has been employed for printing frequency and Eq. (2) has been employed for printed droplet diameter. For the calculation of the composite desirability index 'D' Eq. (3) has been used. The calculated values are given in Table 2. For the calculation of composite desirability index, respective weight values of the response variables should be selected. The weight values have been taken as 0.5 for both the response variable. Equal weight has been selected as both the responses have same importance on the overall printing performance, i.e., getting droplets having smaller diameter as important as achieving greater printing frequency.

Table 2 Individual and composite desirability index

Experiment no	Individual desirability index 'd_i'		Composite desirability 'D'	Rank
	Ave. drop. freq. (Hz)	Ave. drop. dia. (μm)		
1	0.032308721	0.643678161	0.144209633	7
2	1	0.137931034	0.371390676	3
3	0.099776078	1	0.315873516	4
4	0.834385087	0.034482759	0.169622816	6
5	0.126821182	0.459770115	0.241471716	5
6	0.713554541	0.827586207	0.768458129	2
7	0	0	0	8
8	0.620467037	0.988505747	0.78315722	1

The calculated 'd_i' values for all the observations are shown in Table 2. In desirability function analysis the optimal process setting is the one which gives the maximum 'D', i.e., composite desirability value [22–24]. The 'D' values are arranged according to their order in Table 2 under the rank column. From Table 2, it can be seen that observation no. 8 gives the maximum 'D' value. Greater 'D' value implies that, the operating conditions are in favorable state. Therefore, the observation which gives the maximum 'D' value, parameter settings corresponding to that observation gives the optimal result. The optimal values of applied pressure, the standoff height between the nozzle and the substrate, and the applied voltage during printing obtained from proposed methodology are 0.25 psi, 50 μm, and 315 V respectively. Thus the optimization problem has been performed with respect to single composite desirability index which is easy to calculate. Moreover this methodology tries to optimize the multi-response problem which Taguchi S/N ratio analysis is unable to do.

4 Conclusions

In the present study, the desirability function analysis has been proposed as a way of optimization of EHD inkjet-based micro-manufacturing process control parameters. It is a very useful and effective technique for multi-response optimization problem. The optimal combination of the factors is calculated on the basis of composite desirability values, and optimal setting of input process parameters has been identified. Hence, optimization of process parameters with conflicting multi-performance characteristics can be greatly simplified through this approach.

Acknowledgements The authors gratefully acknowledge all the support from Department of Mechanical Engineering, NIT Durgapur.

References

1. Huebner, A., Sharma, S., Srisa-Art, M., Hollfelder, F., Edel, J.B., Demello, A.J.: Microdroplets: a sea of applications? Lab Chip **8**, 1244–1254 (2008). doi:10.1039/b806405a
2. Barton, K., Mishra, S., Alleyne, A., Ferreira, P., Rogers, J.: Control of high-resolution electrohydrodynamic jet printing. Control Eng. Pract. **19**, 1266–1273 (2011). doi:10.1016/j.conengprac.2011.05.009
3. Taylor, G.: Disintegration of water drops in an electric field. Proc. R. Soc. A 280–383 (1964). doi:10.1098/rspa.1964.0151
4. Park, J.U., Hardy, M., Kang, S.J., Barton, K., Adair, K., Mukhopadhyay, D.K., Lee, C.Y., et al.: High-resolution electrohydrodynamic jet printing. Nat. Mater. **6**, 782–789 (2007). doi:10.1038/nmat1974
5. Zeleny, J.: Instability of electrified liquid surfaces. Am. Phys. Soc. **10**, 124–142 (1917). doi:10.1103/PhysRev.10.1

6. Chen, C.H., Saville, D.A., Aksay, I.A.: Electrohydrodynamic "drop-and-place" particle deployment. Appl. Phys. Lett. **88** (2006). doi:10.1063/1.2191733
7. Wang, K., Stark, J.: Direct fabrication of electrically functional microstructures by fully voltage-controlled electrohydrodynamic jet printing of silver nano-ink. Appl. Phys. A **99**, 763–766 (2010). doi:10.1007/s00339-010-5701-5
8. Kim, J.H., Lee, D.Y., Hwang, J., Jung, H.I.: Direct pattern formation of bacterial cells using micro-droplets generated by electrohydrodynamic forces. Microfluid. Nanofluid. **7**, 829–839 (2009). doi:10.1007/s10404-009-0441-6
9. Youn, D., et al.: Electrohydrodynamic micropatterning of silver ink using near-field electrohydrodynamic jet printing with tilted-outlet nozzle. Appl. Phys. A **96**, 933–938 (2009). doi:10.1007/s00339-009-5262-7
10. Jayasinghe, S., Qureshi, Q., Eagles, P.: Electrohydrodynamic jet processing: an advanced electric-field-driven jetting phenomenon for processing living cells. Small **2**, 216–219 (2006). doi:10.1002/smll.200500291
11. Chen, C.H., Saville, D.A., Aksay, I.A.: Scaling laws for pulsed electrohydrodynamic drop formation. Appl. Phys. Lett. **89** (2006). doi:10.1063/1.2356891
12. Choi, H.K., Park, J.U., Park, O., Ferreira, P.M., Georgiadis, J.G., Rogers, J.A.: Scaling laws for jet pulsations associated with high-resolution electrohydrodynamic printing. Appl. Phys. Lett. **92** (2008). doi:10.1063/1.2903700
13. Kim, Y.J., Kim, S.Y., Lee, J.S., Hwang, J., Kim, Y.J.: On-demand electrohydrodynamic jetting with meniscus control by a piezoelectric actuator for ultra-fine patterns. J. Micromech. Microeng. **19** (2009). doi:10.1088/0960-1317/19/10/107001
14. Mishra, S., Barton, K.L., Alleyne, A.G., Ferreira, P.M., Rogers, J.A.: High-speed and drop-on-demand printing with a pulsed electrohydrodynamic jet. J. Micromech. Microeng. **20** (2010). doi:10.1088/0960-1317/20/9/095026
15. Barton, K., Mishra, S., Shorter, K.A., Alleyne, A., Ferreira, P., Rogers, J.: A desktop electrohydrodynamic jet printing system. Mechatronic **20**, 611–616 (2010). doi:10.1016/j.mechatronics.2010.05.004
16. Choi, H.K., Khan, A., Rahman, K., Kwan, K.R.: Effects of nozzles array configuration on cross-talk in multi-nozzle electrohydrodynamic inkjet printing head. J. Electrostat. **69**, 380–387 (2011). doi:10.1016/j.elstat.2011.04.017
17. Naveen Sait, A., Aravindan, S., Noorul Haq, A.: Optimisation of machining parameters of glass-fibre-reinforced plastic (GFRP) pipes by desirability function analysis using Taguchi technique. Int. J. Adv. Manuf. Technol. **43**, 581–589 (2009). doi:10.1007/s00170-008-1731-y
18. Harrington Jr., E.: The desirability function. Ind. Qual. Control **21**, 494–498 (1965)
19. Puviyarasan, M., Senthil Kumar, V.S.: An experimental investigation for multi-response optimization of friction stir process parameters during fabrication of AA6061/B4Cp composites. Arab. J. Sci. Eng. **40**(2015), 1733–1741 (2015). doi:10.1007/s13369-015-1654-5
20. Derringer, G.: Simultaneous optimisation of several response variables. J. Qual. Technol. **12**, 214–219 (1980)
21. Graf, P.G.: A 2nd generation electrohydrodynamic jet (e-jet) printing system, parametric studies of e-jet nozzles and integrated electrode e-jet deposition. M.S. thesis, University of Illinois at Urbana-Champaign. https://www.ideals.illinois.edu/handle/2142/24288 (2011)
22. Sharma, V., Kumar, V.: Multi-objective optimization of laser curve cutting of aluminium metal matrix composites using desirability function approach. J. Braz. Soc. Mech. Sci. Eng. **38**, 1221–1238 (2016). doi:10.1007/s40430-016-0487-9
23. Jayaraman, P., Kumar, L.M.: Multi-response optimization in turning of AA6061 T6 using desirability function analysis. Appl. Mech. Mater. **812**, 124–129 (2015)
24. Kumar, V., Kumar, V., Jangra, K.K.: An experimental analysis and optimization of machining rate and surface characteristics in WEDM of Monel-400 using RSM and desirability approach. J. Ind. Eng. Int. **11**, 297–307 (2015). doi:10.1007/s40092-015-0103-0

A Heuristic Path Search for Congestion Control in WSN

Ganesan Sangeetha, Muthuswamy Vijayalakshmi, Sannasi Ganapathy and Arputharaj Kannan

Abstract There are several factors that affect the performance of a wireless sensor network (WSN) namely storage capacity, energy loss, change in topology, network congestion, deployed environment, intermediate medium used for communication. Out of all these factors of WSN, this paper addresses two major challenges to be refined to decrease the depletion of energy and data loss due to congestion across the network. Initially, node degree and topology of sensor nodes are adjusted periodically at a regular time interval. This ends up in saving the battery power consumption of sensor nodes. The latter factor is a check done for congestion avoidance, by performing rate change using fuzzy logic to balance data flow. Moreover, once when congestion has occurred the search for the best path to reach; sink node is done using LRTA* (Learning Real-Time A Star) heuristic algorithm. Simulations were done to compare the network lifetime of the proposed congestion control mechanism with existing methods. Results show that the heuristic approach for congestion check performs better to the traditional methods.

Keywords Node degree · Wireless sensor network · Congestion control · Fuzzy logic · Rate rectifier · LRTA* algorithm

G. Sangeetha (✉) · M. Vijayalakshmi · A. Kannan
Department of Information Science & Technology, Anna University, Chennai, India
e-mail: sangcse10@gmail.com

M. Vijayalakshmi
e-mail: vijim@annauniv.edu

A. Kannan
e-mail: kannan@annauniv.edu

S. Ganapathy
School of Computing Science and Engineering, VIT University-Chennai Campus, Chennai, India
e-mail: sganapathy@vit.ac.in

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_47

1 Introduction

In wireless sensor networks (WSNs), energy consumption is a major factor that drains the battery power resulting in depleted nodes. Each node in a network can send and receive data, besides forwarding traffic in the network not related to its use. To maintain the location topology and resources, each node has to save its battery power [1]. Energy is largely consumed by increase in the number of unwanted sensor nodes, congestion, unbalanced load state at the nodes, inappropriate topology changes, improper cluster coverage, blind flooding, failed fault tolerant nodes, etc. Proper utilization of battery power would sustain the sensor node in the cluster for further reuse to transfer data packets. The two major classifications of congestion are Content-based Congestion and Buffer-based Congestion. The former occurs when several nodes within the range of one another transmit at the same time resulting in data loss due to interference. The latter occurs when each node uses a buffer for data packets to be sent. Overflow in buffer causes congestion and packet loss. Several parameters are used to indicate congestion in sensor nodes namely packet loss, queue length, channel load, packet service time, delay, scheduling time.

Congestion Control is an another important factor that has to be considered next to energy efficiency with respect to the QoS (Quality of Service) requirements of WSNs. However, by inspecting and avoiding congestion, the consumption of battery power is considerably reduced. The limitations of a wireless sensor network lossy data communication, adaptability into a new environment, processing battery power, reframing topology each time and resource storage limit. A wireless sensor network (WSN) consists of a set of tiny nodes, accommodated with embedded computing devices communicable with sensors/actuators. They are independently acting wireless transmitters that are used within a short range. WSNs route data packets by hops to reach the common collecting node called sink node or base station.

There are several techniques to perform congestion control in WSNs to minimize packet loss rate supporting battery power reduction. A priority-based congestion control mechanism for WSNs using Fuzzy Logic is given to adjust the level of data packet transfer, once when congestion occurs. The offspring sensor nodes receive a notification to reduce their transmit rate using a Fuzzy Rate Adjustment Inference System (FRAIS) [2]. This method differentiates locally generated data and transit data using a priority allocation thereby providing an efficient queuing model. Proposed FRAIS in PCCF (Predictive Congestion Control scheme using Fuzzy Logic) model [2] shows that there is considerable decrease in packet drop rate and end-to-end delay when compared to Congestion Control Scheme based on Fuzzy logic theory (CCSF). In this paper, first a periodic timer is set to reform the clusters each time periodically. In the latter phase, a congestion control mechanism with fuzzy rule set is combined with prioritized LRTA* algorithm to service the sensor node in path selection.

The rest of the paper is structured as follows: Sect. 2 presents the related discussions to the contributed work. Section 3 represents system model. Description

about Congestion Checkpoint and Congestion Control mechanism is given in Sect. 4. Section 5 explains about the simulation results obtained using MATLAB. The proposed work is concluded with future directions in Sect. 6.

2 Related Work

Several explorations have been done to minimize congestion and to boost up the battery power efficiently in WSNs. Earlier congestion control techniques included balancing load conditions on the sensor nodes. Energy efficiency can be achieved in WSNs by predicting congestion control in the network. A probabilistic means is proposed to predict congestion using network traffic and buffer size. Rate allocation schemes [3] such as rate reduction (RR), rate regulation (RRG) and split protocol (SP) are listed to improve network performance and reduce packet loss rate. This technique outperforms the decentralized predictive congestion control method (DPCC) [4].

There are many traffic control protocols [1] implemented in WSNs. The first classification includes Equal or Weighted Partitioning. Apart from a list of protocols available, there are two protocols namely CCF (Congestion Control and Fairness) [5] and Prioritizing Information for QoS control. The former protocol is a many-to-one routing protocol to modify data rate using the packet service time which is the period between sending data packet at transport layer to network layer and receiving successful transmitted notification. The latter protocol gives priority to sensed data based on its features. This uses end-to-end packet delay to sense congestion and report data rate at sink. Clustering is a strong approach to arrange sensor nodes in an associated order to provide balanced load distribution across nodes. It also increases the battery power lifetime of WSN. A cluster head close to the sink node consumes its energy quickly developing hot spot problems. To overcome this issue, many algorithms on unequal clustering have been proposed. Among them, a Fuzzy Based Unequal Clustering [6] is presented to improve energy efficiency of WSN. Fuzzy rule sets are given for node degree and computing competition radius for the members to join cluster head. Many cluster based routing algorithms have been proposed by many researchers for energy-efficient routing in WSN [7, 8]. This approach leads to saving a considerable amount of battery power during heavy congestion in the network.

A Traffic-Aware Dynamic Routing (TADR) algorithm [9] is proposed to perform routing of packets around heavy congestion areas. In addition, excessive packets are distributed along multiple paths comprising idle and underloaded sensor nodes. A hybrid virtual potential field is created to guide the packets from obstacles and to move towards the sink node. Sink mobility faces new challenges to be addressed namely topology framework, load estimation, connectivity, data storage, and network coverage. Adaptive routing combined with load estimation [10] technique can be used for sink relocations. This method replaces periodic traffic injection of less path trace quality that over utilizes the network resources.

The Learning Real-Time Algorithm (LRTA*) [11] is a heuristic algorithm which was developed according to [12], guided by other heuristic algorithms like Enforced Hill Climbing (EHC) and Best First Search (BFS). LRTA* possess the advantage of performing optimization such as pruning successors during expansion. It also maintains a heap with maximum buffer capacity to store the visiting states during the search. For mission-critical applications, LRTA* would be suitable as it has a polynomial complexity across a state space.

3 System Model

The proposed system assumes the presence of one sink node and a set of sensor nodes within each cluster C that is defined as follows:

$$C_n = [SN_1, SN_2, SN_3, \ldots, SN_i] \quad (1)$$

where, n indicates the number of nodes in a single cluster. A static distribution of clusters is considered across the widespread geographical area. Each of the cluster has a collection of scattered nodes within it. A non-moving environment of the sensor node clusters is considered. The sensor nodes are randomly placed across the area and all nodes contained within a cluster are homogeneous that are identical with same storage space, performance, computation power, architecture, and processing. The structure of the network comprises of a chain of nodes connected to form a cluster, and in turn a chain of clusters to form the wireless sensor network. There is a Cluster Head (CH) elected to coordinate the sensor nodes within it.

Each sensor node has different battery power levels. The selection of CH in a cluster depends on two main factors: residual energy of sensor nodes and node degree framed between the nodes. Ultimately the alteration of node degree each time gives uniformly load balanced clusters, minimizing energy consumption and congestion for every cycle. Figure 1 shows the place where congestion takes place.

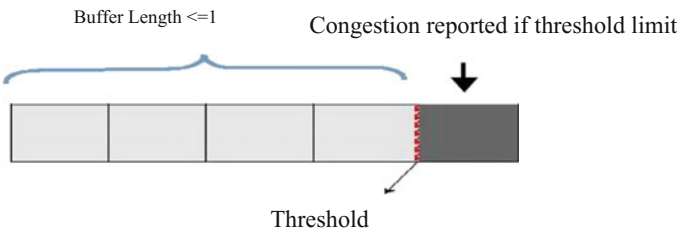


Fig. 1 Point of congestion occurrence

4 Congestion Checkpoint and Control Mechanism

This section describes about congestion avoidance and the remedial action to be taken once congestion has been detected in the sensor network system.

4.1 Congestion Factor

Congestion Factor (CF) gives a numerical factor value for congestion at each node n and is calculated for all successor nodes. The corresponding calculation for CF is defined as follows:

$$CF = (1 - rate^{n}_{incoming} / rate^{n}_{outgoing}) \text{ x } T \quad (2)$$

where, rate $_{incoming}^{n}$ represents incoming packet traffic rate at the current node n and rate $_{outgoing}^{n}$ represents outgoing packet traffic rate at current node n. Further, rate $_{incoming}^{n}$ can be defined as follows:

$$rate^{n}_{incoming} = r^{n}_{s} + r^{n}_{tr} \quad (3)$$

where, r_s^n represents source traffic rate, locally generated data from source node n and r_{tr}^{i} represents transit traffic, the data transmitted using intermediate nodes at n. So a congestion is said to not occur when congestion factor lies intermediate between 0 and 1 ($0 \leq CF\ 1$). A congestion between intermediate nodes occur when congestion factor is less than 0 ($CF < 0$).

4.2 Buffer Length

Buffer Length (BL) assumes two sensor nodes A and B. A buffer length variable depends on the incoming packet and assistance rate at each neighboring node. The neighbor buffer length state can be visualized by a pair of ways such as piggy-backing and beacon broadcast.

$$BL(i) = P^{i}_{QB} / P^{i}_{TB} \quad (4)$$

where P_{QB}^{i} represents the number of packets in queue buffer at node i and P_{TB}^{i} represents the number of packets in total buffer size at node i. The buffer length is in the range 0 <= BL(i) <= 1. Buffer Length exceeds threshold value if BL(i) = 1.

4.3 Proposed Framework

The proposed system combines the incoming fuzzifier output into a Fuzzy Rate Rectifier System (FRRS). A Heuristic Path Selection (HPS) with a goal-based path search algorithm LRTA* is used. We propose a model of Heuristic-based Priority Congestion Control scheme with Fuzzy logic theory (HPCCF) in Fig. 2 to trace congestion and perform heuristic path selection. The modules designed address three main workings

(i) Periodic refreshment of CH selection by changing the cluster formation each time.
(ii) When congestion is found, intimated to their clusters and incoming data rate is adjusted.
(iii) At once congestion is come across from then choose only the best path using LRTA* goal reaching to sink node.

The rate of adjustment is indicated by means of a notification signal to rectify and adjust transmission rate using Fuzzy Rate Rectifier System (FRRS). There are

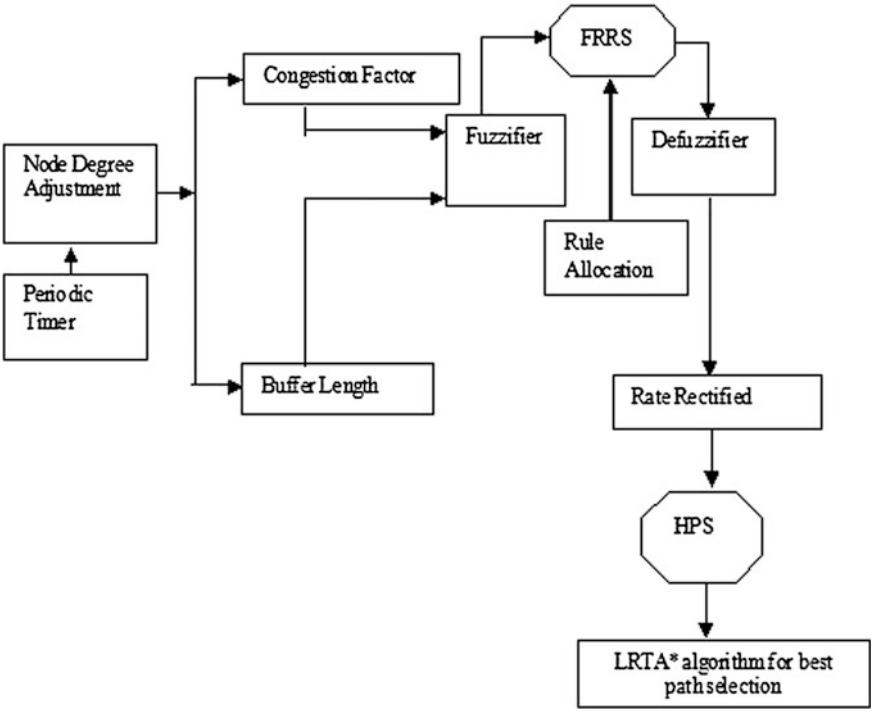


Fig. 2 Proposed fuzzy system with heuristic path selection

two types of triggering about congestion using signal: Implicit Congestion Control (ICN) and Explicit Congestion Control (ECN). In former signaling, notification is sent along with the data packet. In latter, notification is sent as separate packet.

In the proposed method, ICN is used to send notification signal to all neighboring nodes. ICN is used due to its ability to avoid heavy traffic blocking across the network compared to ECN. However, ECN performs notification without dropping packets.

4.3.1 Fuzzification Process

Fuzzy logic is used as a mode for incoming data rate adjustment. Fuzzy sets allow the data members to be partially located or associated to a set. Assume A_1, A_2 and A_3 are crisp sets. Then γ_{B1}: $A_1 -> [0, 1]$, γ_{B2}: $A_2 -> [0, 1]$ and γ_{B3}: $A_3 -> [0, 1]$ are termed

as the membership functions of B_1, B_2 and B_3 that define the fuzzy sets B_1, B_2, B_3 of A_1, A_2, A_3. Fuzzification process maps two crisp sets with their parameters as shown below

B_1 = *Congestion Factor* ϵ *{low, medium, high}*
B_2 = *Buffer Length* ϵ *{low, medium, high}*

The membership function B_3 after applying input parameters to Fuzzy Rate Rectifier System (FRRS) is obtained as

B_3 = *Rate Rectified* ϵ *{DVL, DL, DM, DH, DVH}*

To extract the fuzzy system output based on the inputs applied to the system, fuzzy rules are defined. So far there are nine fuzzy rules, since there are two fuzzy sets each consisting of three states. The total number of possible fuzzy rules is 9 ($3^2 = 9$).

4.3.2 Best Goal-Based Path Selection

Once congestion has occurred across sensor network, by default selecting exact appropriate path by hops to sink node is activated. This is done by using LRTA* search algorithm. The main purpose of selecting the best path is to reduce heavy congestion thereby saving energy consumption in the network.

The output of B_3 (Rate Rectifier for Congestion) represents the rate rectification due to congestion and consists of five fuzzy states: DVL (Decrease Very Low), DL (Decrease Low), DM (Decrease Medium), DH (Decrease High), and DVH (Decrease Very High). B_4 output (HPS Priority) represents the priority states taken by LRTA* algorithm to treat the most appropriate path to sink node thereby saving battery power consumption. All fuzzy values of B3 and B4 have been normalized between [0,1]. Membership function B_4 notifies the action to be taken as a result of congestion occurrence.

Consider the Path Selection Point to be final output fuzzy set B_4. When congestion has occurred in the system, fuzzy set B_4 value is given taking the best path point score from LRTA* algorithm. Applying inputs from rate rectified B_3 to Heuristic Path Selector (HPS) the final output fuzzy set for best path to sink node is given as below,

Path Selection Point (PSP) ϵ {normal, intermediate, immediate}

B4 = {VLN, LN, INT, IMM, VIMM}

A trapezoidal function to show the fuzzy sets. The trapezoidal function T is defined as below with four parameters

$$T(b_1, a, m_2, \beta) = \begin{cases} \frac{b_1 - \alpha}{m_1 - \alpha} & if\ b_1 \in [\alpha, m_1] \\ 1 & if\ b_1 \epsilon\ [m_1, m_2] \\ \frac{\beta - b_1}{\beta - m_2} & if\ b_1 \epsilon\ [m_2, \beta] \\ 0 & Otherwise \end{cases} \tag{5}$$

In this computation, α and β are the valleys and m_1 and m_2 are maximum points of the trapezoidal function. This trapezoidal function mapping b_1 lies in between [0, 1] and degree of membership called $\gamma(b_1)$ has been generated. To obtain the membership degrees of B_1, B_2 and B_3 sets we define relations. For degree B_1 computation is given as

$$\begin{aligned} \text{low} &= \text{trap}(\text{w: } 0, 0, \text{m}/2, \text{m}) \\ \text{medium} &= \text{trap}(\text{w: } \text{m}/2, \text{m}, \text{m}, \text{m}/2 + \text{m}) \\ \text{high} &= \text{trap}(\text{w: } \text{m}, \text{m} + \text{m}/2, \text{infinity}, \text{infinity}) \end{aligned} \tag{6}$$

The above relations can be used for computing B_2 and B_3. For B_4, consider the overlook B_1, B_2 and B_3 to give the priority decision B_4 to reach the sink node using goal search algorithm.

4.3.3 Defuzzification Process

AND logic is to write fuzzy rules, to obtain $\gamma(b_3)$ and $\gamma(b_4)$, the minimum value between $\{\gamma(b_1)$ and $\gamma(b_2)\}$ and $\{\gamma(b_1)$, $\gamma(b_2)$ and $\gamma(b_3)\}$ is found. Minimal value is computed as,

$$\gamma_{b1 \cap b2} = \min(\gamma_{b1}, \gamma_{b2}) \tag{7}$$

$$\gamma_{b1 \cap b2 \cap b3} = \min(\gamma_{b1}, \gamma_{b2}, \gamma_{b3}) \tag{8}$$

The mathematical result of the fuzzy system, the crisp output b_3 and b_4, center of sums (COS) from defuzzification method is used. In this technique, the fuzzy logic controller calculates the geometric center of area (COA) for every membership function as below:

$$\text{COA (FRRS)} = \frac{\int \mu_{b3}(b_3)\, b_3\, d_{b3}}{\int \mu_{b3}(b_3).\, d_{b3}} \tag{9}$$

$$\text{COA (HPS)} = \frac{\int \mu_{b4}(b_4)\, b_4\, d_{b4}}{\int \mu_{b4}(b_4).\, d_{b4}} \tag{10}$$

Fuzzy Logic Controller calculates the area below the already scaled membership functions and within the range of output variable. The above equations are used to calculate the geometric center of this area.

5 Simulation Results and Discussions

The performance of the proposed technique is compared with the earlier scheme CCSF and the successor scheme PCCSF [2] in terms of energy consumption and congestion level. The simulations of the proposed method are performed using MATLAB. The number of nodes in simulated network is 500 and they are randomly deployed in 50 × 50 m^2 field where are there are five source nodes and one sink node. The available buffer size of each node is 50 packets and packet length is fixed at 250 bytes. Assumption is made that all nodes are static with no mobility in WSN.

Figure 3 illustrates the total energy consumed in CCSF, PCCF and for the proposed method. The proposed method is able to utilize less battery power due to two major reasons. First, node clustering is done each time to elect the longer lasting node to withstand data transmission to sink overcoming congestion in WSN. The latter is the use of a heuristic search algorithm to select the exact path to sink, without any wastage of energy level. These solutions enable HPCCF to consume less battery power when compared to the above-existing methods. For every

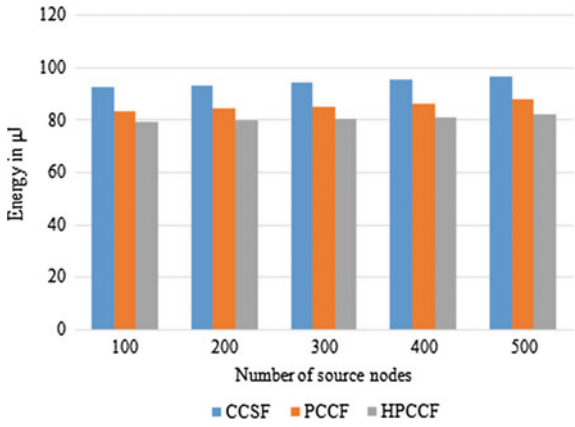


Fig. 3 Comparison in terms of energy consumption

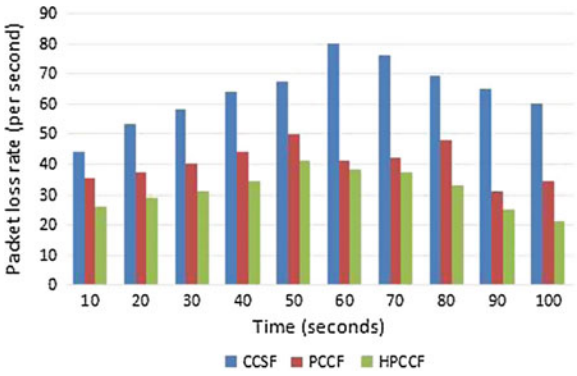


Fig. 4 Comparison in terms of packet loss due based on congestion level

collective 100 nodes, the proposed method HPCCF utilizes lesser energy to carry on data transmission in WSN.

Figure 4 shows the total packet loss rate for the existing methods and proposed HPCCF method. The packet drop rate is considerably lesser in the proposed technique when compared to CCSF and PCCF. For example, up to 30 s there is only a minor packet drop rate, which increases and then drops down again resulting in a minimal packet loss rate.

From Figs. 3 and 4, it can be observed that the performance of the proposed system is better when it is compared with the existing systems. This is due to the fact that the use of fuzzy rules with heuristic algorithm.

6 Conclusion

In this paper, the main QoS factors in wireless sensor networks such as reduction energy consumption and congestion control has been achieved. Initially the power utilized was reduced using repeated periodic cluster formation to select new CH that can withstand efficiently. Congestion is a major factor that has to be considered to decrease power consumed in WSN. When congestion occurs, the best path selection is done using heuristic LRTA* search algorithm, that is a mission critical to spot the goal. Fuzzy logic rules are used to take the priority decision to be treated with search algorithm to reach sink node. This saves the energy usage by checking congestion in WSN. Simulation results show that the proposed work is better than CCFL in terms of energy loss rate and packet loss rate. So, the proposed framework HPCCF minimizes end-to-end delay by reducing data loss and time required to reach sink node. The future direction of this work is to have a comparison of other goal-based search algorithms to find the best path to sink node and to integrate them into usage in congestion control.

References

1. Kafi, M.A., Djenouri, D., Othman, J.B., Badache, N.: Congestion control protocols in wireless sensor networks: a survey. IEEE Commun. Surv. Tutorials **16**, 1369–1390 (2014)
2. Hatamain, M., Barati, H.: Priority—based congestion control mechanism for wireless sensor networks using fuzzy logic. In: IEEE 6th ICCNT, July 2015, pp. 13–15 (2015)
3. Rajan, A.U., Raja, S.V., Jeyasekar, A., Lattanze, A.J.: Energy-efficient predictive congestion control for wireless sensor networks. IET Wirel. Sens. Syst. **5**(3), 115–123 (2015)
4. Zawodniok, M., Jagannathan, S.: Predictive congestion control protocol for wireless sensor networks. IEEE Trans. Wireless Commun. **6**(11), 3955–3963 (2007)
5. Yin, X., Zhou, X., Huang, R., Fang, Y., Li, S.: A fairness-aware congestion control scheme in wireless sensor networks. IEEE Trans. Veh. Technol. **58**(9), 5225–5234 (2009)
6. Logambigai, R., Kannan, A.: Fuzzy logic based unequal clustering for wireless sensor networks. Elsevier. Int. J Wirel. Netw. **22**(3), 945–957 (2016)
7. Jothi, M.S., Kannan, A., Ganapathy, S.: Virtual force based clustering for energy efficient routing in mobile wireless sensor networks. Transylvanian Rev. **24**(9), 1–8 (2016)
8. Logambigai, R., Ganapathy, S., Kannan, A.: Cluster based routing with isolated nodes in WSN. Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET) **4**(3), 343–34 (2016)
9. Fengyuan, R., Tao, H., Sajal, K.D., Chuang, L.: Traffic-aware dynamic routing to alleviate congestion in wireless sensor networks. IEEE Trans. Parallel Distrib. Syst. **22**(9), 1585–1599 (2011)
10. Karenos, K., Kalogeraki, V.: Traffic management in sensor networks with a mobile sink. IEEE Trans. Parallel Distrib. Syst. **21**(10), 1515–1530 (2010)
11. Alves, R.M.F., Lopes, C.R., Branquinho, A.A.B.: Generating plans using LRTA*. In: Brazilian Conference on Intelligent Systems, pp. 207–212 (2013)
12. Rios, L.H.O., Chaimowicz, L.: A survey and classification of A* based best-first heuristic search algorithms. In: Advances in Artificial Intelligence, 20th Brazilian Symposium on Artificial Intelligence Proceedings, pp. 253–262 (2010)

Alignment-Independent Sequence Analysis Based on Interval Distribution: Application to Subtyping and Classification of Viral Sequences

Uddalak Mitra and Balaram Bhattacharyya

Abstract Viral sequence classification has widespread applications in structural and functional categorization, clinical and epidemiological studies. Most approaches of subtyping and classification depends on an initial alignment step to asses similarity score followed by distance-based phylogenetic or statistical algorithms. We observe that interval distributions of nucleotide(s) over the sequence possess the potential for sequence comparison and devise an algorithm that determines the similarity/dissimilarity score among pairs of sequences. Classification of HIV virus subtyping by the method obtains exact tally with its biological taxonomy.

Keywords Interval distribution · HIV subtyping · Pattern based classification · k-word pattern in HIV genomes

1 Introduction

Genome sequencing has reached a new era with advent of new generation sequencing (NGS) technologies resulting voluminous amount of molecular sequence data. Bioinformatics tools, techniques, and approaches have begun to decipher functional and structural features using sequence data with objective to link genotypes and phenotypes, study of molecular evolution and facilitate the task of taxonomic classification, reorganization of genera, families, and subfamilies.

Both subtyping and classification require assessment of sequence similarity in their workflow. Alignment-dependent similarity assessment methods suffer by quadric time complexity and alignment quality is hard to relate with statistical relevance of similarity score [1, 2]. To overcome the aforesaid limitations, efforts are on to evolve alignment-independent methods. Approaches based on alignment-independent assessment of similarity try to extract inherent sequence features and compare them

U. Mitra · B. Bhattacharyya (✉)
Department of Computer and System Sciences, Visva-Bharati University, Santiniketan 731235, India
e-mail: balaramb@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_48

in a statistically meaningful way. Methods based on k-words frequency are predominant of alignment-independent methods.

K-word based feature frequency profile (FFP) [3, 4] has been successfully applied to devise similarity in whole-genome phylogeny. It computes the counts of each possible k-word in given sequences. Each k-word count in each sequence is then divided by total k-word counts that generate a numeric vector of size 4^k for a given k value. JS- divergence between two numeric vectors is used as a distance measure between two genomic sequences. Distance matrix computed from each possible pair-wise distance is then used to derive phylogenetic relationship among the genome sequences. Composition vector method [5, 6] basically follows the steps of FFP method, except it subtract the random background from frequency count of each k-word using a Markov model of order (k-2) that diminish the influence of random neutral mutations. However, all these approaches try to use only frequency information and differences in genome lengths generate unequal number of observations in terms of frequency count, thus showing inability to represent complete pattern of information content in sequences.

In addition to the frequencies of k-words, subsequent studies [7, 8] put additional stress on their locational information and order of occurrence along the sequence. In DMk [7], a numeric vector of size 4^k, for a given k, summarizes the whole sequence. Euclidian distance between numeric vectors of two genomes is used as dissimilarity measure. The Category Position Frequency model [8] is an improved version of DMk, where each genome sequence is at first converted into three unique representations of symbols based on chemical property of nucleotides. Then from each symbolic representation, the method generates four numeric values using similar technique of DMk to construct the numeric vector that represent the given sequence. Both the methods work efficiently with coding regions, but do not corroborate with biological taxonomy when applied to whole-genome similarity assessment of HIV genome subtyping and classification. Furthermore, the similarity scores of both the methods do not possess upper and lower limits, making it difficult to relate with statistical relevance. In addition, there is no one-to-one mapping between computational scores and actual similarity.

An in-depth study on occurrences of successive k-words reveals that information on intervals between successive occurrences of optimal k-words contains potential intrinsic features of the sequence that is crucial for comparison between a pair of sequences. The present work attempts to explore these features.

2 Problem Identification and Our Contribution

Identification of HIV virus subtypes has impact on disease progression, response in treatment and vaccine design and development. Virus subtyping and classification of types or subtypes is also informative and a reliable tool in epidemiology and surveillance, design of preventions and transmission of follow-up. Yet identification

of correct HIV subtypes is a challenging task due to unique recombinant forms (URFs).

In this paper, we propose a method, which is nearly passive to recombination effect and utilizes the information inherent in interval distribution of nucleotide(s) of a given sequence to form a feature vector used as representative of the sequence. Finally, we develop a dissimilarity matrix using Bhattacharyya distance [9] to study the degree of dissimilarity between pairs of sequences using feature vector representations.

3 Materials and Method

3.1 Exploring Patterns

A genome sequence is composed of the repeated occurrences of the four nucleotides {A, T, G and C}. An interval is a gap between successive appearances of a particular nucleotide or a particular group of nucleotides, termed as k-word. Successive occurrences of different k-words over the stretch of the sequence provide information about structure of the sequence and we observe that the distribution of the intervals among the successive k-words provides an important clue for comparison of sequences. We begin to unfold it with the example in Fig. 1.

All the intervals of 'A' in the sample sequence in Fig. 1 is simply computed by subtracting the current location of appearance from the next appearance except the last interval; for example, the first interval 2 of 'A' is computed as $8 - 6 = 2$; here 6 is the current location of appearance and 8 is the next. The last interval value '8' for 'A' is chosen by assuming circular relationship between last and first appearance of 'A'. In the sample sequence last 'A' appear at location 38 and first 'A' appear at location 6, thus we calculate the interval between the two appearances of 'A' as $40 - 38 + 6 = 8$, here 40 is the length of the sample sequence. Similar intervals can be drawn for T, C, G and for the different combinations of nucleotides. How much informative are the intervals that we illustrate with real dataset in the following section.

Initially we check the intervals for nucleotide adenine (A) of Human Immunodeficiency Virus-1 (*accession no. AB098330*), Simian Immunodeficiency Virus (SIV) (*accession no. AB808655*), Hepatitis C Virus (HCV) (*accession no. AB016785*) and Hemorrhagic Fever Virus (HFV) (*accession no. NC_016144*) and

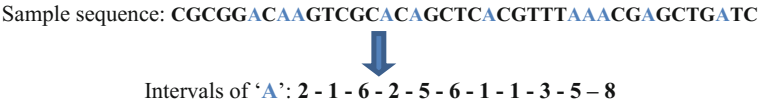


Fig. 1 Computation of intervals of 'A' in the sample sequence at k = 1

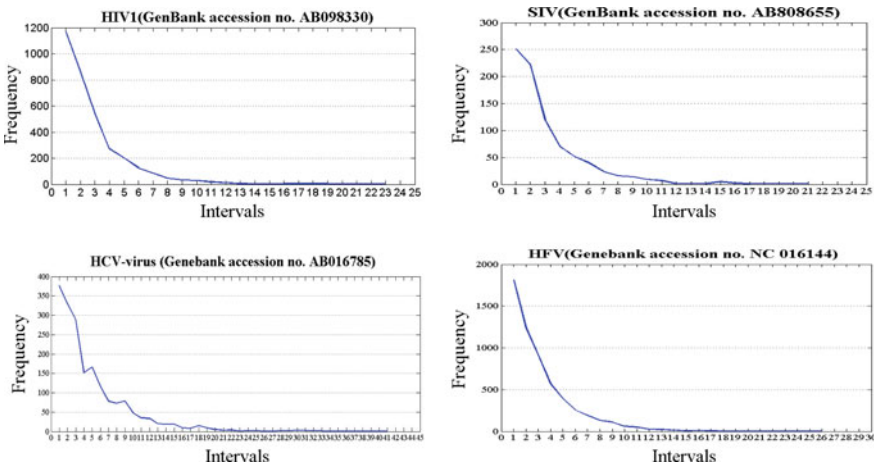


Fig. 2 The frequency plots for interval distributions of Adenine (A; k = 1) in virus genomes

summarized in the frequency plot of Fig. 2. Variation as measured in terms of interval distribution of 'A' of these viruses that belongs to different virus family reflected in these plots (*HIV and SIV belongs to family: Retroviridae, genus: Lentivirus, HCV belongs to family: Flaviviridae, genus: Hepacivirus and HFV belongs to family: Filoviridae, genus: Cuevavirus*). For example, the frequencies of interval value '1' of nucleotide 'A' in case of HIV-1 virus is 1166, for SIV it is 251, for HCV it is 377 and that for HFV it is 1818. Moreover, all interval distributions are positively skewed and the pattern of peaks is showing local variations. In additional, the plot also displays lower intervals are highly frequent than higher intervals and this is due to local dependency property of nucleotide sequences. We also observed variations in terms of interval distribution of other nucleotides(C, G, T) in the genomes of these viruses (not shown in Figure). Therefore, the interval distribution of four nucleotides, at k value 1 together can summarize the genomic contents with inherent property to recognize patterns.

Similar analysis was carried out with four virus genomes viz. HIV-1 virus subtype A1 (Genebank accession no. AB098331), HIV-1 virus subtype B (Genebank accession no. AB078005), HIV virus subtype C (Genebank accession no. AB023804), and HIV 1 virus subtype D (Genebank accession no. A07108). It can be clearly observed that, the extent of variations in the patterns of interval distributions for viruses within a genus Fig. 3 is less than that in genera (Fig. 2) as expected. Therefore, sequence similarity can be assessed by homogeneous patterns in interval distributions and serves as the basis of our proposed method.

Frequency plot can primarily serve as a tool to determine similarity among virus genomes but differences in lengths generates unequal number of observations in terms of intervals. Thus, it is essential to represent the interval distributions with statistical parameters and we use Information Theoretic entropy(h) [10] to summarized an interval distribution. Procedure for computation of entropy to

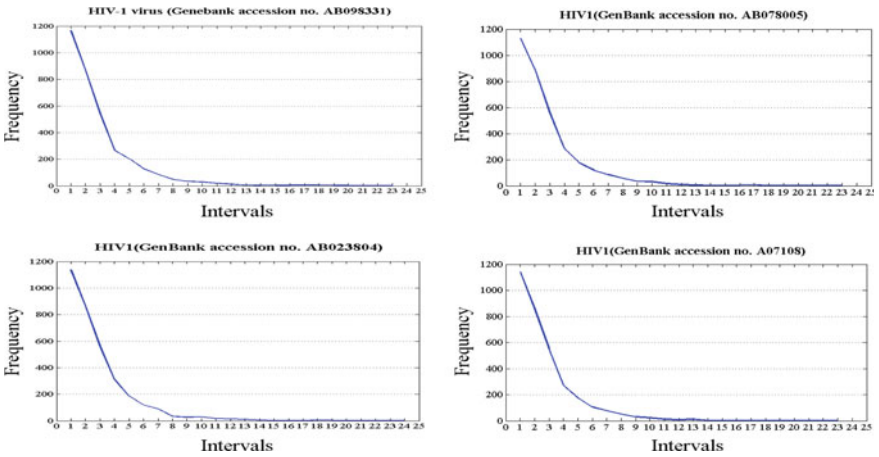


Fig. 3 The frequency plots for interval distributions of Adenine (A; $k = 1$) in HIV virus

Interval distribution of 'A'

Intervals of 'A'	Probability mass function
1	3*1/40
2	2*2/40
3	1*3/40
5	2*5/40
6	2*6/40
8	1*8/40

Entropy of interval distribution of 'A': h = (-3 * (1/40) * l(2/40) * log (2 / 40)) + (- 1 * (3/40) * log(3/40)) + (- 2 * (5/40) * log(5/40)) + (- 2 * (6 / 40) * log(6/40)) + (- 1 * (8/40) * log(8/40) = 1.9215

Fig. 4 Computation of interval distribution of 'A' and its statistical parameter h in the sample sequence at k = 1

summarized interval distribution of nucleotide 'A' for the sample sequence is exemplified in Fig. 4. The statistical summary of the interval distributions of nucleotides A, T, C and G using entropy(h) is shown in Table 1 for the viruses discussed above.

Since a nucleotide sequence is composed of four bases, adenine (A), guanine (G), thymine (T), and cytosine (C), for a subsequence of nucleotides of size k, theoretically, 4^k k-words are possible and thus 4^k corresponding interval distribution to each k-word for a sequence can be computed. We thus proceed to extract patterns from interval distributions of k-words, at a given value of k, over the sequence which is formalized in the following section.

Table 1 Statistical summary of interval distributions of the representative virus genomes

Species	Entropy of interval distribution of A	Entropy of interval distribution of C	Entropy of interval distribution of G	Entropy of interval distribution of T
HIV-1virus (AB098330)	0.2655	0.2389	0.2497	0.2459
SIV virus (AB808655)	0.2622	0.2376	0.2492	0.2510
HCV-virus (AB016785	0.2423	0.2582	0.2542	0.2453
HFV-virus (NC_016144)	0.2584	0.2515	0.2432	0.2468
HIV-1 virus (AB098331)	0.2655	0.2387	0.2498	0.2460
HIV-1 virus (AB078005	0.2651	0.2385	0.2505	0.2458
HIV-1 virus (AB023804)	0.2652	0.2380	0.2500	0.2467
HIV-1 virus (A07108)	0.2666	0.2381	0.2494	0.2459

3.2 Formalizing the Method

Let $\sum = \{A, C, G, T\}$ be an alphabet and S is a genome sequence constructed by succession of 'n' symbols from $\sum$. A fragment of 'k' ($\leq n$) of consecutive symbols in S is designated as a k-word. With four symbols in $\sum$, the set of k-words for a particular k becomes $\{w_1, w_2, \ldots, w_r\}$ (where $r = 4^k$). Let us denote the jth k-word by w_j, that occurs at locations l_i^j, i = 1(1)m, m being the maximum number of occurrence. We define a sets of intervals between each successive locations of that jth k-word as $D\{d_1^j, d_2^j, \ldots, d_m^j\}$: where

$$\begin{aligned} d_i^j &= l_{i+1}^j - l_i^j \text{ for } i < m \\ &= |S| - l_m^j + l_1^j \text{ for } i = m \end{aligned} \tag{1}$$

An obvious feature of D is that sum of its elements equals the length of the sequence

$$\sum_{i=1}^{m} di = |S| \tag{2}$$

If $p_i = d_i / \sum_{i=1}^{m} d_i$, i = 1(1)m, be the probability of occurring the jth k-word, w_j at intervals d_i^j, then the statistical parameter h_j, represent the interval distribution of the jth k-word, defined according to Shannon's entropy as

$$h_j = \sum_{i=1}^{|D|} p_i * \log(1/p_i) \tag{3}$$

Entropies of all the k-words form a numeric vector of size 4^k. Upon normalization, we obtain the representation vector of the given sequence S for word size k

$$H = (h_1, h_2, \ldots, h_r)./ \sum_{j=1}^{r} h_j \tag{4}$$

where ./ denotes element-wise division.

H is a numeric vector of size $r = 4^k$, each element represents normalized entropy of the interval distribution of a k-word of the sequence at a given k value.

3.3 Distance Measure Between Genomes Based on Parameter of Interval Distributions

To measure the distance between two genome sequences (S_i and S_j) we use Bhattacharyya distance [9] between given pair of genome sequences, as follows:

$$d_{ij} = -\log\left(\sum \left(H_{irh} * H_{jrh}\right)^{1/2}\right) \tag{5}$$

'r' stands for interval distribution of a particular k-word. The terms H_{irh} and H_{jrh} means normalized entropy of rth interval distribution in ith genome and jth genome respectively, where $r = 1, \ldots, 4^k$. In case of the absence of a k-word or single occurrence, we assign zero to the parameter for that k-word.

3.4 Workflow of the Method

Input: Sequences $\{S_1, S_2, \ldots, S_N, k\}$

Output: Distance matrix d(N × N): each element holds pair-wise similarity

Steps:

1 For each sequence, search and locate each k-word of word size k
1.1 For each k-word, use Eq. (1) to calculate intervals
1.2 For each k-word, use Eqs. (2) and (3) to calculate entropy of the interval distribution
2. For each sequence, construct 4^k component numeric vector using Eq. (4)
3. For any two sequences, use Eq. (5) to calculate distance between the two sequence
4. Return {d}.

3.5 Time Complexity of the Algorithm

The proposed algorithm makes feature vector of a DNA sequence in $O(n + 4^k \times 4m)$ time, where n is the length of the DNA sequence, k is the optimal word length and m is the average number occurrence of a word in the sequence. If each word is equal-probable to occur, the average word count can be given by $m = n/4^k$. Thus making feature vector requires time $O(n + 4^k \times 4m)$, that is, $O(n + 4^k \times 4n/4^k) = O(5n)$. Thus the algorithm possesses linear time complexity.

4 Experiment

The series of experiments are conducted on laptop with Intel i5 processor with 4 GB RAM using MATLAB environment. The datasets are taken from freely available repository of Los Alamos National Laboratory, USA (http://www.hiv.lanl.gov/content/sequence/HIV/mainpage.html).

4.1 Datasets: HIV Virus Genomes

Human Immunodeficiency Virus one (HIV1) appears as a number of distinct types, subtypes, and recombinant forms. HIV1, the predominant stain of worldwide epidemic is divided into six groups (HIV-1 M group, HIV-1 N group, HIV-1 O groups, HIV-1 P group, HIV-2, and SIV). M group is further divided into nine subgroups (A, B, C, D, F, G, H, J and K) and subgroup A and F are further divided into sub-subgroups (A1, A2) and (F1, F2) respectively.

4.1.1 Training Data

Dataset A consists of 55 representative genomes of HIV1 virus repository of Los Alamos National Laboratory with known genotypes. This dataset is used as training dataset for HIV subtyping

4.1.2 Classification Data

Dataset B consists of 75 HIV virus genomes downloaded from the same source and there is no common genome in A and B, i.e., $A \cap B = \Phi$. The dataset B is used as a true positive dataset to test the performance of the method for classification of HIV viruses. Details of the data sets are listed in supplementary file (on request from author)

4.2 Results and Discussion

The interval distribution provides a technique to represent a genome sequence in a compact manner and is computationally tractable in straightforward way. Once given genomic sequences are converted in the form of numeric vectors it now becomes computationally efficient to compute pair-wise genome distances using Eq. 5. Sequences having similarities are grouped together, exhibiting homogeneous patterns of interval distributions, based on their k-word dependencies.

4.2.1 Subtyping of HIV Virus

Human Immunodeficiency Virus belongs to the family of Retroviridae, genus Lentivirus that causes HIV-infection and may lead to Acquired Immunodeficiency Syndrome (AIDS). AIDS in human causes progressive failure of the immune system resulting life-threatening opportunistic infections and cancers to thrive.

One of the obstacles of standard treatment of HIV/AIDS is high genetic variability as it is very common that a therapy fruitful for a particular genetic stain becomes irreverent for another. Thus, a fast identification of subtypes is a need.

Applicability of the proposed method for subtyping was tested using HIV genome sequences of dataset A (supplementary Table ST1) to train the system. Cladogram trees were inferred at different k values (range from 1 to 8) using both the N-J method and the UPGMA method. The accuracy of clustering was assessed based on patterns of clades being generated in the cladograms. We observe that k value 6 is sufficient to correctly classify the subtypes of HIV viruses in dataset A. Figure 5 shows the cladograms of the dataset A. It is clear from the figure that the intra-subtypes are grouped together and inter-subtypes are placed in distant clades in the tree. The result is also 100% consistent with the taxonomic classification provided in LANL (summarized in Table 2). Trees inferred using DMk [7] and CPF [8] are given in supplementary Figure SF2 and SF3 reflect the inconsistency of the results found using these methods.

4.2.2 Classification of HIV Viruses

It was seen that the proposed method has correctly clustered the HIV genotypes of dataset A at k value 6. In order to benchmark the proposed method in terms of performance in classification of HIV viruses, the true positive dataset B (supplementary table ST2) was used. For each of these genome sequences in dataset B, interval distributions of all k-words at k value = 6 was calculated using the workflow shown in Sect. 3.2 and then distance was calculated using Eq. 5 from the representative genomes of dataset A. Based on the idea of closeness of

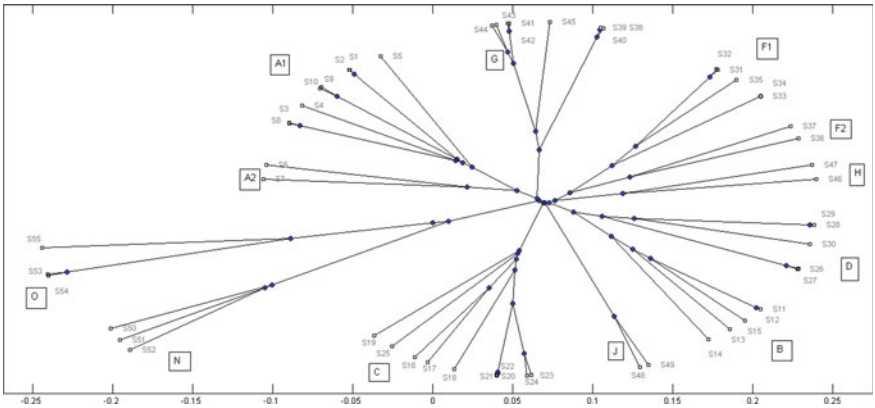


Fig. 5 Cladogram obtained using genomes of HIV viruses of dataset A using the proposed method at k value 6 and clustering method Neighbor-Joining. Letters and associated numerals in square boxes are showing the HIV subtypes and sub-subtypes. *Note* Cladogram using UPGMA method was also tested and result is consistent with the shown figure

Table 2 HIV subtypes of dataset A with their respective labels according to LANL

Label	S1 to S5	S6 to S7	S8 to S10	S11 to S15	S16 to S25	S26 to S30	S31 to S35	S36 to S37	S38 to S45	S46 to S47	S48 to S49	S50 to S52	S53 to S55
Subtype	A1	A2	A1	B	C	D	F1	F2	G	H	J	N	O

distance, we select the subtype of the reference genome (in dataset A), which has minimum distance with the given query genomes (in dataset B), as the predicted subtype for that query genome. The query genomes (Qi, 1 = 1(1)75) with their predicted subtypes and respective minimum distance from the selected reference genome of training dataset A is shown in Table 3. It was observed that all the query genome sequences in true positive dataset B were correctly assigned to their genotypes. The result of classification when compared with the COMET [11] tools for HIV subtyping generates some unclassified stains are found and marked by (*). We also evaluate ranges of distances within the dataset A and B that were estimated to derive cut-off values. The empirically devised cut-offs is used to eliminate possible false positive predictions by the method. A true negative dataset consists of 47 HIV genomes that was derived from dataset A that excludes subtype A1. These true negative dataset and dataset B are used to devise sensitivity and specificity of the method. It is observed that the method works with 100% sensitivity and 100% specificity.

Table 3 Classification of HIV subtypes of dataset B with their respective distances

Label	Subtypes	Distance	Label	Subtypes	Distance	Label	Subtypes	Distance
Q1	A1-1	0.1016	Q26	C-16	0.0824	Q51	G-45	0.1185
Q2	A1-1	0.1016	Q27	C-16	0.0761	Q52*	G-45	0.1187
Q3*	A1-2	0.1095	Q28	C-16	0.0687	Q53	G-45	0.1339
Q4*	A1-3	0.1091	Q29	C-16	0.0575	Q54	H-47	0.1095
Q5	A1-1	0.0998	Q30*	C-16	0.0726	Q55	H-46	0.1063
Q6	A1-1	0.0995	Q31	D-30	0.1084	Q56	J-48	0.1265
Q7	A1-1	0.1085	Q32	D-30	0.1020	Q57	N-51	0.0948
Q8	A1-9	0.1108	Q33	D-30	0.1004	Q58	N-51	0.1035
Q9	A1-4	0.1082	Q34	D-30	0.1130	Q59	O-53	0.1519
Q10	A1-5	0.1134	Q35	D-29	0.1151	Q60	O-53	0.1259
Q11	A1-10	0.1130	Q36	D-30	0.1523	Q61	O-53	0.0880
Q12	A1-1	0.1671	Q37	D-26	0.1503	Q62	O-53	0.1135
Q13	B-15	0.1108	Q38	D-26	0.1455	Q63	O-53	0.1095
Q14	B-15	0.1108	Q39	D-26	0.1776	Q64	O-53	0.0750
Q15	B-15	0.0001	Q40	D-26	0.1770	Q65	O-53	0.1442
Q16	B-15	0.1050	Q41	F1-31	0.1010	Q66	A1-5	0.1376
Q17	B-15	0.1042	Q42	F1-32	0.1090	Q67	A1-5	0.1356
Q18	B-15	0.1057	Q43	F1-33	0.1211	Q68	A1-4	0.1219
Q19	B-15	0.1185	Q44	F1-31	0.0459	Q69	A1-4	0.1228
Q20	B-15	0.1186	Q45	F1-31	0.0173	Q70	A1-4	0.1195
Q21	C-22	0.0043	Q46	F1-35	0.1169	Q71	A1-4	0.1194
Q22	C-22	0.0049	Q47	F1-33	0.1080	Q72	A1-4	0.1082
Q23	C-20	0.0044	Q48*	F1-33	0.1190	Q73	A1-4	0.1117
Q24	C-22	0.0052	Q49	F2-36	0.1011	Q74	A1-5	0.1106
Q25	C-22	0.0076	Q50*	F2-26	0.1214	Q75	A1-5	0.1097

5 Conclusion

In the proposed method, we develop an alignment-independent dissimilarity measure between two nucleotide sequences. We have applied the distance measure in subtyping and classification analysis of HIV virus genomes to assess applicability of the method. On subtype analysis, the method constructs trees that are found to be 100% consistent with biological taxonomy provided in LANL. The result of classification when compared with COMET tools shows improved accuracy with 100% sensitivity and specificity.

References

1. Vinga, S., Almeida, J.: Alignment-free sequence comparison-a review. Bioinformatics (Oxford Journal) **19**(4), 513–523 (2002)
2. Bonham-Carter, O., Steele, J., Bastola, D.: Alignment-free genetic sequence comparisons: a review of recent approaches by word analysis. Bioinformatics (Oxford Journal) **15**(6), 890–905 (2013)
3. Sims, G.E., Jun, S.R., Wu, G.A., Kim, S.H.: Whole-genome phylogeny of mammals: evolutionary information in genic and nongenic regions. Proc. Natl. Acad. Sci. U.S.A. **106** (40), 17077–17082 (2009)
4. Sims, G.E., Kim, S.H.: Whole-genome phylogeny of Escherichia coli/Shigella group by feature frequency profiles (FFPs). Proc. Natl. Acad. Sci. U.S.A. **108**(20), 8329–8334 (2011)
5. Gao, L., Qi, J.: Whole genome molecular phylogeny of large dsDNA viruses using composition vector method. BMC Evol. Biol. (2007)
6. Wang, H., Xu, Z., Gao, L., Hao, B.: A fungal phylogeny based on 82 complete genomes using the composition vector method. BMC Evol. Biol. **9**, 195 (2009)
7. Wei, D., Jiang, Q., Wei, Y., Wang, S.: A novel hierarchical clustering algorithm for gene sequences. BMC Bioinform. **13**(174) (2012)
8. Bao, J., Yuan, R., Bao, Z.: An improved alignment-free model for dna sequence similarity matric. BMC Bioinform. 15(321) (2014)
9. Bhattacharyya, A.: On a measure of divergence between two statistical populations defined by their probability distributions. Calcutta Math. Soc. **35**, 99–109 (1943)
10. Shannon, C.E.: A mathematical theory of communication. Bell Syst. Tech. J. **27**(3), 379–423 (1948)
11. Struck, D., Lawyer, G., Ternes, A.M., Schmit, J.C., Perez Bercoff, D.: COMET: adaptive context-based modeling for ultrafast HIV-1 subtype identification. Nucleic Acids Res. **42**, e144 (2014)

Part IX
Biotechnology, Biomedical Instrumentation and Bioinformatics

Automated Glaucoma Detection from Fundus Images of Eye Using Statistical Feature Extraction Methods and Support Vector Machine Classification

Abhishek Dey and Kashi Nath Dey

Abstract Glaucoma is one of the eye diseases that can lead to the blindness if not detected and treated at proper time. This paper presents a novel technique to diagnose glaucoma using digital fundus images. In this proposed method, the objective is to apply image processing and machine-learning techniques on the digital fundus images of the eye for separating glaucomatous eye from normal eye. Image preprocessing, techniques such as noise removal and contrast enhancement are used for improving the quality of image thus making it suitable for further processing. Statistical feature extraction methods such as Gray-Level Run Length Matrix (GLRLM) and Gray-Level Co-occurrence Matrix (GLCM) are used for extracting texture features from preprocessed fundus images. Support Vector Machine (SVM) classification method is used for distinguishing glaucomatous eye fundus images from normal, unaffected eye fundus images. The performance of the trained SVM classifier is also tested on a test set of eye fundus images and comparison is done with other existing recent methods of Glaucoma detection.

Keywords Glaucoma · Fundus · Feature extraction · Support vector machine

1 Introduction

Glaucoma is a disease which affects optic nerve and it can result in progressive, permanent visual loss. The eye is filled with aqueous, a fluid, which is there at a stable pressure called intraocular pressure (IOP). Generally, glaucoma is

A. Dey (✉)
Department of Computer Science, Bethune College, Kolkata, India
e-mail: dey.abhishek7@gmail.com

K.N. Dey
Department of Computer Science and Engineering, University of Calcutta, Kolkata, India
e-mail: kndey55@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_49

developed due to an elevation in the IOP, although additional factors can play roles in its development. In some cases, glaucoma may occur in the presence of normal eye pressure. Poor regulation of blood flow to the optic nerve can be another cause of glaucoma. The damage to the optic nerve initially results in loss of peripheral visual fields and later it affects the central vision as well. Glaucoma is the second leading cause of blindness in the world and its progression is expected to increase [1].

Early detection and subsequent treatment is necessary for affected patients to preserve their vision because lost capabilities of the optic nerve cannot be recovered. Manual analysis of the eye is very much time consuming and the accuracy of the parameter measurements also varies with different clinicians. The cost of detecting glaucoma using OCT and HRT is very high. Alternatively fundus cameras can be used by many ophthalmologists to diagnose glaucoma. Images processing techniques allow us to extract features from fundus images of eye that can provide useful information to diagnose glaucoma [2].

Several techniques have been used to automate the glaucoma detection process in recent times. In [1], glaucoma image classification using fuzzy min-max neural network based on Data-Core (DCFMN) has been proposed. An automated method for detecting glaucoma from color fundus images is proposed in [2] where after preprocessing techniques, an appearance-based dimension reduction process is used to compress different generic features. These features are combined by a probabilistic two-stage classification scheme to extract the novel Glaucoma Risk Index (GRI). In [3], cup-to-disc ratio of optic disc of eye is determined using Harris Corner that measures the local changes of the signal with patches shifted in different directions by a small amount. [4] presents a novel method for glaucoma detection from fundus images using Haralick Texture Features and K-Nearest Neighbors (KNN) classifiers. [5] presents glaucoma detection using Artificial Neural Network (GD-ANN) where neuron model has been developed using feed-forward back-propagation network. The parameters used in this method are Intra Ocular Pressure, Central Corneal Thickness, Nerve Fiber Layer Thickness, and CDR.A novel method for detecting glaucoma has been proposed in [6], based on principle components analysis (PCA) and Bayes classifier.

This paper presents a novel method for glaucoma detection from eye fundus images using image processing and machine-learning techniques. The goal of the system is to separate glaucomatous eye fundus images from a set of eye fundus images given as input to the system. The rest of the paper is organized as follows. Section 2 elaborates the proposed methodology that includes image preprocessing techniques, texture feature extraction using statistical methods and classification using Support Vector Machine. Experimental results are shown in Sect. 3. Finally, the conclusion is summarized in Sect. 4.

2 Methodology

At first, the images of eye fundus are preprocessed. Then, several texture features are extracted from preprocessed images by using statistical methods Gray-Level Co-occurrence Matrix (GLCM) and Gray-Level Run Length Matrix (GLRLM). The extracted features are fed into a Support Vector Machine (SVM) classifier in the form of a feature matrix. The classifier is trained by using K-fold cross validation approach. A new set of eye fundus images are tested by the trained classifier for performance evaluation. The block diagram of the proposed system is shown in Fig. 1.

2.1 Preprocessing

Preprocessing can drastically improve the performance of image processing methods like feature extraction, segmentation and disease detection. Preprocessing eliminates disease independent variations from the input image, suppresses undesired distortions or enhances some image features, thus making it suitable for subsequent processing tasks [7]. The steps involved in preprocessing are diagrammatically presented in Fig. 2. The original raw input images are normalized in the range [0, 1]. After color conversion, the grayscale images are resized to 256 × 256. Then, two-dimensional Gaussian Filter is used for noise removal. After

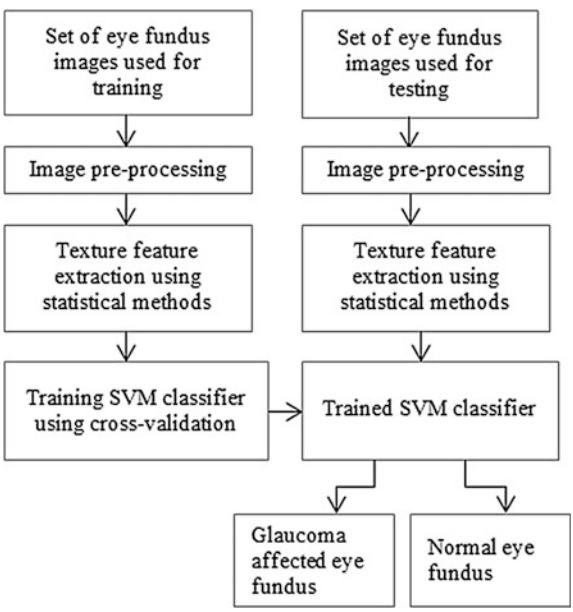


Fig. 1 Block diagram of the proposed system

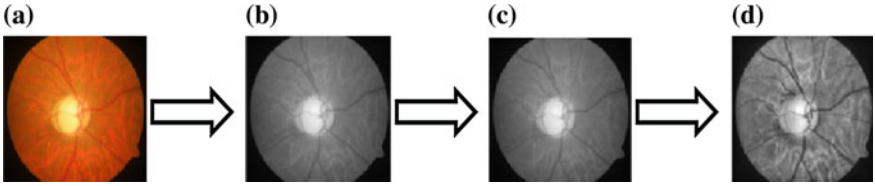


Fig. 2 **a** Input image before preprocessing. **b** Eye fundus after Normalization and color conversion. **c** Eye fundus after noise removal. **d** Preprocessed eye fundus after contrast enhancement

noise removal, Adaptive Histogram Equalization technique is used for improving the image contrast. This step also removes the nonuniform background which may be due to nonuniform illumination or variation in the pigment color of eye.

2.2 Feature Extraction

The objective of feature extraction is to reduce the original data by measuring positive properties, or features, that discriminate one input sample from another sample. The texture of images refers to the structure, arrangement, and appearance of the parts of an object within the image [8]. In this paper, 16 texture features are extracted from preprocessed input images by using statistical methods as explained in the remaining part of this section.

Entropy is extracted directly from preprocessed grayscale image. It is defined as $H = -\sum p \log_2 p$, gives a statistical measure of randomness that can be used to characterize the texture of the input image, where p contains the histogram counts of the image [9]. Gray-Level Co-occurrence Matrix (GLCM) describes the spatial relationship between each intensity tone by considering changes between gray levels i and j at a particular distance d and at a particular angle θ. Formally, the $G \times G$ Gray-Level Co-occurrence Matrix P for a displacement vector $d = (dx, dy)$ is defined as follows. The entry (i, j) of P is the number of occurrences of the pair of gray levels i and j which are at distance d apart. Formally, $P(i, j) = \{((r, s), (t, v)): I(r, s) = i, I(t, v) = j \mid (r, s), (t, v) \in N \times N, (t, v) = (r + dx, s + dy)\}$ [10]. After creating GLCM of a preprocessed input image, we have used $d = 1$ and $\theta = 0°$ for extracting following four features from the matrix:

Contrast:

$$\sum_{ij} |i - j|^2 P(i, j), \tag{1}$$

measuring of the intensity contrast between a pixel and its neighbor over the whole image.

Correlation:

$$\sum_{i,j} \frac{(i-\mu i)(j-\mu j)P(i,j)}{\sigma_i \sigma_j}, \tag{2}$$

measuring how a pixel is correlated to its neighbor over the whole image, where μi, μj denotes mean values of i and j respectively and σ_i, σ_j denotes standard deviations of i and j respectively

Energy:

$$\sum_{i,j} P(i,j)^2, \tag{3}$$

measuring sum of squared elements in the GLCM

Homogeneity:

$$\sum_{i,j} \frac{P(i,j)}{1+|i-j|}, \tag{4}$$

measuring the closeness of the distribution of elements in the GLCM to the GLCM diagonal

In run length matrix $P_\theta(i, j)$, each cell in the matrix consists of the number of elements where gray level i successively appears j times in the direction θ. j is termed as run length [11]. The resultant matrix characterizes gray-level runs by the gray tone, length, and the direction of the run. This method allows extracting the higher order statistical texture features. Texture properties obtained from GLRLM include the following five main features:

Short run emphasis:

$$\frac{\sum_i \sum_j \frac{p_\theta(i,j)}{j^2}}{\sum_i \sum_j p_\theta(i,j)}, \tag{5}$$

measuring fine texture

Long run emphasis:

$$\frac{\sum_i \sum_j j^2 P_{\theta^{(i,j)}}}{\sum_i \sum_j P_{\theta^{(i,j)}}} \tag{6}$$

measuring coarse texture

Run length nonuniformity:

$$\frac{\sum_j \sum_i \left\{P_{\theta^{(i,j)}}\right\}^2}{\sum_i \sum_j P_{\theta^{(i,j)}}}, \tag{7}$$

measuring similarity of the run lengths

Gray-level nonuniformity:

$$\frac{\sum_i \sum_j \left\{P_{\theta(i,j)}\right\}^2}{\sum_i \sum_j P_{\theta(i,j)}}, \tag{8}$$

measuring similarity of gray-level values

Run percentage:

$$\frac{\sum_i \sum_j P_{\theta(i,j)}}{A}, \tag{9}$$

measuring the number of short runs, where A is the area of interest in the image

These fivefeatures along with sixmore features such as low gray-level run emphasis, high gray-level run emphasis, short run low gray-level emphasis, short run high gray-level emphasis, long run low gray-level emphasis, and long run high gray-level emphasis are extracted from horizontal direction ($\theta = 0°$) of GLRLM of preprocessed images.

2.3 Classification

A classification task separates data into training and testing sets. Each training set instance contains one target value (i.e., the class labels) and several attributes (i.e. the features). We have used Support Vector Machine (SVM) classifier, a supervised learning model, for classifying normal eye fundus from glaucoma affected eye fundus. SVM is designed to separate a set of training images into two different classes, (x_1, y_1), (x_2, y_2), … (x_n, y_n) where $x_i \in R^d$, the d-dimensional feature space, and $y_i \in \{-1, +1\}$, the class label, with $i = 1 \ldots n$. SVM builds the optimal separating hyper planes based on a kernel function (K). All images, of which feature vector lies on one side of the hyper plane, belong to class −1 and the others are belong to class +1.

Simplest type of SVM is linear SVM, in which the training patterns are linearly separable. A linear function of this form is shown below

$$f(x) = w^T x + b \tag{10}$$

such that for each training sample x_i the function yields $f(x_i) > 0$ for $y_i = +1$, and $f(x_i) < 0$ for $y_i = -1$. In other words, training samples of two different classes are separated by the hyperplane $f(x) = w^T x + b = 0$, where w is weight vector and normal to hyperplane, b is bias or threshold and x_i is the data point. In Fig. 3, linear SVM classification with a hyperplane that minimizes the separating margin between the two classes are indicated by data points by white and black circles. Support

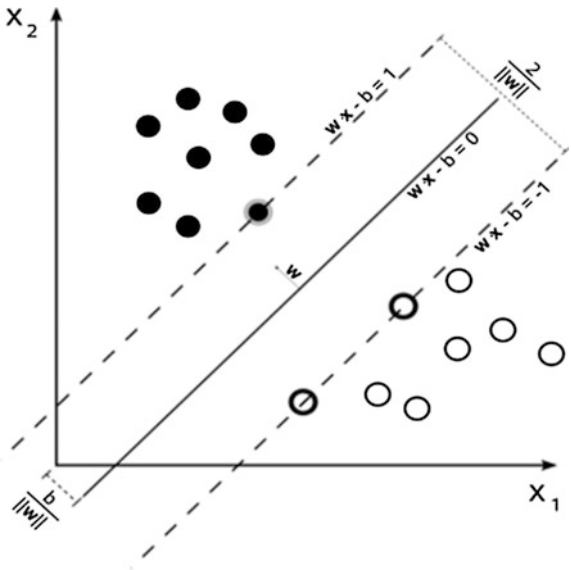


Fig. 3 Maximum-margin hyper plane and margins for a linear SVM trained with samples from two classes

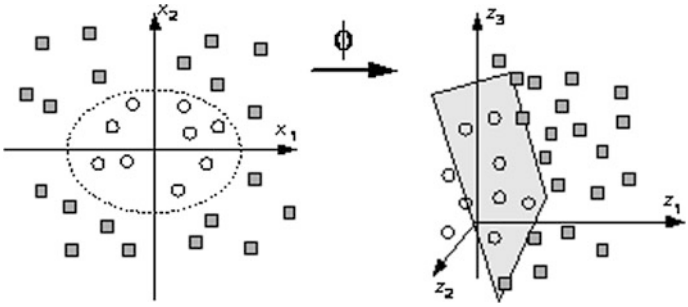


Fig. 4 Nonlinear SVM classification

vectors are elements of the training set that lie on the boundary hyperplane of the two classes [9].

If the data of the classes cannot be separated, then the nonlinear SVM classifier is used [9]. In a nonlinear SVM classifier, a nonlinear operator is used to map the input pattern x into a higher dimensional space H as shown in Fig. 4.

The nonlinear SVM classifier is defined as,

$$f(x) = w^T \phi(x) + b \tag{11}$$

The transformation from nonlinear to linear separating hyperplane in higher dimensional feature space is done by taking help of kernel functions. A kernel function on two samples, represented as feature vectors in some input space, is defined as $k(x_i, x_j) = \phi(x_i)^T \phi(x_j)$, ϕ is the feature vector. Most commonly used kernels are

$$\text{Linear kernel: } k\left(x_i, x_j\right) = x_i^T x_j \tag{12}$$

$$\text{Polynomial Kernel: } k\left(x_i, x_j\right) = \left(\gamma x_i^T x_j + r\right)^d, \gamma > 0 \tag{13}$$

$$\text{RBF Kernel: } k\left(x_i, x_j\right) = \exp\left(-\frac{||x_i - x_j^2||}{2\sigma^2}\right), \sigma > 0 \tag{14}$$

$$\text{Sigmoid Kernel: } k\left(x_i, x_j\right) = \tan h\left(\gamma x_i^T x_j + r\right) \tag{15}$$

σ, γ, r, and d are all kernel parameters.

RBF kernel generally performs better than other kernel functions when number of observations is larger than number of features [12]. Proper selection of kernel parameters is necessary for achieving very high accuracy of classification. In our proposed method, we have compared performances of the above-mentioned kernels for training the SVM classifier.

3 Experimental Results and Discussion

All the image preprocessing, feature extraction, and SVM classification techniques in our proposed method are simulated in MATLAB 8.1 (R2013a) and run on a Intel (R) Core(TM) Quad CPU-Q 8200 PC with 8-GB memory. We have obtained both glaucomatous and healthy eye fundus images with age group of 15 to 70 years from Susrut Eye Foundation and Research Centre, Kolkata. These are all real-world images verified by eye specialists. 100 such images are used as input to our proposed system. This set of 100 images contains 50 glaucoma affected eye and 50 normal eye fundus images. After preprocessing steps, 16 texture features are extracted from these images as described in Sect. 2 and fed into SVM classifier for training purpose. K-fold cross validation method is used for training SVM classifier. We have taken K = 10. Hence, 90 images are used for training and 10 images are used for testing performance each time. This process is repeated 10 times using different sessions of the test images each time. Performance of the classifier can be tested and evaluated by the following parameters:

Accuracy rate = Correctly classified images/Classified images, Sensitivity = Correctly classified positive images/True positive images, Specificity = Correctly classified negative images/True negative images, Positive Predictive Accuracy (PPA) = Correctly classified positive images/Positive classified images, Negative

Table 1 Performance comparison of kernel functions used in training the SVM classifier

Kernel function	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPA (%)	NPA (%)
Linear	86	94	78	81.03	92.86
Polynomial	92	96	88	88.89	95.65
RBF	97	96	98	97.96	96.08
Sigmoid	67	72	62	65.45	68.89

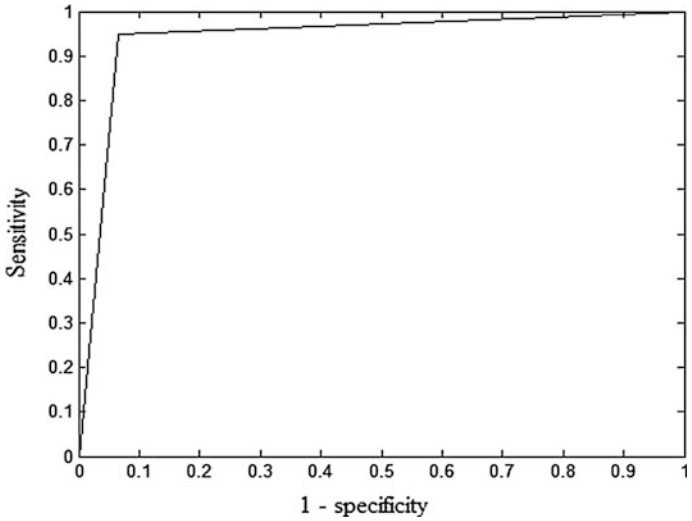


Fig. 5 ROC curve obtained by applying the trained SVM classifier on the test set of images

Predictive Accuracy (NPA) = Correctly classified negative images/Negative classified images.

In our work, SVM classifier is trained using different kernel functions. It is observed from Table 1 that using Radial Basis Function (RBF) kernel (with $\sigma = 1.2$), SVM classifier can achieve highest accuracy of 97% after training. So, we have used the cross-validated trained classifier using RBF kernel for further testing on new fundus images. Performance testing of this trained classifier is done on 50 eye fundus images (20 normal and 30 glaucoma affected) which were not in the set of input images used for training. The SVM classifier can successfully classify this test set with accuracy of 94%, sensitivity 93.33%, specificity 95%, PPA 96.55% and NPA 90.48%.

Figure 5 shows the Receiver operating characteristic (ROC) curve obtained by applying the trained SVM classifier on the test set of fundus images. ROC curve is a performance evaluation curve for binary classifier systems [13]. It is a plot of the true positive rate or sensitivity against the false positive rate or (1—specificity) for different thresholds of the classifier output of a diagnostic test. The area under the

Table 2 Performance comparison of existing methods

Method	Classification technique used	No. of images used	Training accuracy (%)
DCFMN	Fuzzy min-max neural network	40	85
GD-ANN	Artificial neural network	39	97
SVMGD	SVM (with RBF kernel)	100	97

ROC curve (AUC) is a measure of how good a test can distinguish between two diagnostic groups (diseased/normal). An area of 1 represents a perfect test; an area of 0.5 represents a very bad test. In our method, the obtained area of ROC curve is 0.9417 that signifies our method of glaucoma diagnosis is excellent. Comparison of our proposed method with two other existing recent methods DCFMN [1] and GD-ANN [5] is also done in this paper. Our method is named as SVMGD. It can be observed from Table 2, our method SVMGD is optimal than DCFMN and GD-ANN with respect to number of images and also accuracy.

4 Conclusion

In this paper, we have used SVM classifier to develop a system which can automate classification of glaucomatous eye fundus images from healthy eye fundus. The computational overhead of SVM is much less than other classifiers as only a small training set required during training. The SVM classifier is trained using well known kernel functions. RBF kernel outperforms all other kernels and highest accuracy of trained classifier thus achieved is 97%. Also, comparison of our method with other existing methods establishes superiority of our method. Hence, it can be concluded that our proposed method is very much efficient for automated detection of glaucoma from eye fundus images and also computationally very fast. This method is intended to help the ophthalmologists and doctors in their decision making process by freeing them from the burden of time-consuming manual checking of eye fundus images.

References

1. SriAbirami, S., Grace Shoba, S.J.: Glaucoma images classification using fuzzy min-max neural network based on data-core. Int. J. Sci. Mod. Eng. (IJISME), **1**(7) (2013). ISSN: 2319-6386
2. Bock, R., Meier, J., Nyúl, L.G., Hornegger, J., Michelson, G.: Glaucoma risk index: automated glaucoma detection from color fundus images. Med. Image Anal. **14**(3), 471–481 (2010)
3. Dey, N.: Optical cup to disc ratio measurement for glaucoma diagnosis using harris corner. In: ICCCNT12

4. Simonthomas, S., Thulasi, N.: Automated diagnosis of glaucoma using Haralick texture features. IOSR J. Comput. Eng. (IOSR-JCE) **15**(1), 12–17 (2013). e-ISSN: 2278-0661, p-ISSN: 2278-8727
5. Sheeba, O., George, J., Rajin, P.K., Thomas, N., George, S.: Glaucoma detection using artificial neural network. IACSIT Int. J. Eng. Technol. **6**(2) (2014)
6. Xiong, L., Li, H., Zheng, Y.: Automatic detection of glaucoma in retinal images. In: IEEE 9th Conference on Industrial Electronics and Applications, pp. 1016–1019 (2014)
7. Rathinam, S., Selvarajan, S.: Comparison of image preprocessing techniques on fundus images for early diagnosis of glaucoma. Int. J. Sci. Eng. Res. **4**(12) (2013). ISSN 2229-5518
8. Castellano, G., Bonilha, L., Li, L.M., Cendes, F.: Texture analysis of medical images. Clin. Radiol. 1061–1069 (2004)
9. HariBabuNandpuru, D.S.: Salankar, Prof. V.R. Bora; IEEE Students' Conference on Electrical, Electronics and Computer Science, pp. 1–6 (2014)
10. Eleyan, A., Demirel, H.: Co-occurrence matrix and its statistical features as a new approach for face recognition. Turk. J. Electr. Eng. Comput. Sci. **19**(1), 97–107 (2011)
11. Danny, L.Y.T., Udyavara, D.R.A.: Computer Based Diagnosis of Glaucoma Using Principal Component Analysis (Pca): A Comparative Study. SIM University (2011)
12. Hsu, C.W., Chang, C.C., Lin, C.J.: A Practical Guide to Support Vector Classification. Department of Computer Science National Taiwan University, Taipei 106, Taiwan
13. Kumar, R., Indrayan, A.: Receiver operating characteristic (ROC) curve for medical researchers. Indian Pediatr. **48**, 277–287 (2011)

Classification of EEG Analysis for Sleep Detection Using Wavelet Transform and Neural Network

G.K. Rajini, Shaik Naseera, Shikha Pandey and Akshaya Bhuvaneswaran

Abstract This paper develops a drowsiness detecting system by resolving the Electroencephalogram (EEG) signals. The acquired EEG signals are subjected to noise and are removed by subtracting the artifacts from the original EEG recording using Biorthogonal Wavelet Filter. Features are extracted using Discrete Wavelet Transform in particular Daubechies' wavelet with five-level decomposition is utilized to segregate the signal into five sub-bands, namely, delta, theta, alpha, beta and gamma. The statistical moments such as mean, variance, standard deviation and power of the signal are calculated and stored. These moments serve as an input to the next stage, i.e., system classification. Unsupervised learning using K-means clustering is employed as the classification of the signals are not known. Following this, Support Vector Machine and Pattern Recognition Network are employed for supervised classification. This system provides strong decision making during a real-time sleep detection scenario.

Keywords EEG-electroencephalograph signals · DWT-discrete wavelet transform · LBP-local binary pattern · SVM-support vector machine · PRN-pattern recognition network

1 Introduction

In this paper, we aim to implement the structure that classifies EEG signals into respective sleep related classes. Drowsiness poses a major threat to driver's life and innocent pedestrians which can leads to accidents. The positioning of electrodes for gathering EEG is done using a standard system called the 10–20 montage [1].

G.K. Rajini (✉) · S. Pandey · A. Bhuvaneswaran
School of Electrical Engineering, VIT University, Vellore, India
e-mail: rajini.gk@vit.ac.in

S. Naseera
School of Computing Science and Engineering, VIT University, Vellore, India
e-mail: shaik.naseera@vit.ac.in

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_50

The EEG signals collected would be analyzed by DWT that employs Daubechies' wavelet and five-level decomposition. We apply Discrete Fourier Transform of the signals to identify the spectral components of the EEG Signal. The mean, variance, standard deviation and power would be calculated for these coefficients and stored as a set of features. Then, the Local binary pattern is computed to calculate the histogram. Later the EEG signal is subjected to Haar wavelet to extract features which are subsequently used for K-means clustering as the classes of the EEG signal are unknown. Here classification algorithms like Support Vector Machines and Pattern Recognition Network are employed to classify the following 5 cases i.e., awake, drowsy, asleep, head down and eye blink. The system we have proposed aims to calculate 100% accuracy, sensitivity and specificity. EEG signals from Physionet sleep-EDF database [2] has been used to test the algorithm.

2 Related Work

A comprehensive review of study on several issues relating to fatigue detection, image processing, EEG analysis, DWT, Discrete Fourier Transform and classification for solving issues related to drowsiness detection are presented [3]. Peters et al. [3] presented the idea of moving left or right index finger, or right foot which is recognized in electroencephalograms (EEGs). The efficiency observed in the process is 75–98% where speed and reliability proved to be influential. Dong et al. [4] proposed in the paper, driver lethargy detection, which is divided into two categories: (1) distraction and (2) fatigue. It can be classified using the following measures: (1) subjective report measures (2) driver biological measures (3) driver physical measures (4) driving performance measures and (5) hybrid measures. The first two are not suitable for real time applications. Kanungo et al. proposed a notable algorithm for k-means clustering. Results proved that the computational period had been reduced by an order or two.

3 Design Approach and Details

The EEG signal is collected to find the drowsiness of the driver while travelling. Discrete Wavelet Transform (DWT) is employed to analyze the filtered signal. This method uses a mother wavelet for measuring the alikeness between the basis function and our signal in terms of the frequency content [5]. Five sub-bands of the EEG signal will be obtained after decomposition of Daubechies Wavelet. The mean, median, variance, standard deviation and mode from each of the sub-bands are calculated and stored as a set of features. K-Means clustering is employed so as to classify the signals of unknown classes. This is further classified via a three layer artificial neural network into drowsy, awake or sleepy, head down or blinking. Figure 1 represents the flow of the system.

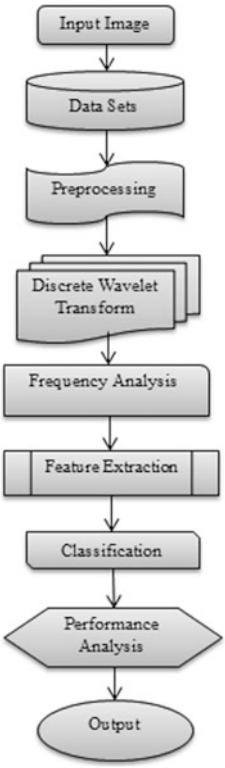


Fig. 1 Flow chart of sleep detection system

Figure 2 illustrates the manner in which various entities have been integrated into one system, and the relationship between these entities which form a vital part of signal processing. They have been further divided into sub-parts which constitute them. It shows how the entities are related to each other and interact among themselves.

4 Description and Goals

Drowsiness leads to accidents by slowing down the decision capability of the driver, causing vision impairment and leading to major faults in judgment [6, 7]. Hence in order to gain better road safety, the following goals will be accomplished in this paper.

(1) To classify the sleep related state of the driver using his EEG signal, and to consider the non-stationary nature of the signal.
(2) To intimate the driver by setting up an alarm system (text and voice intimation).

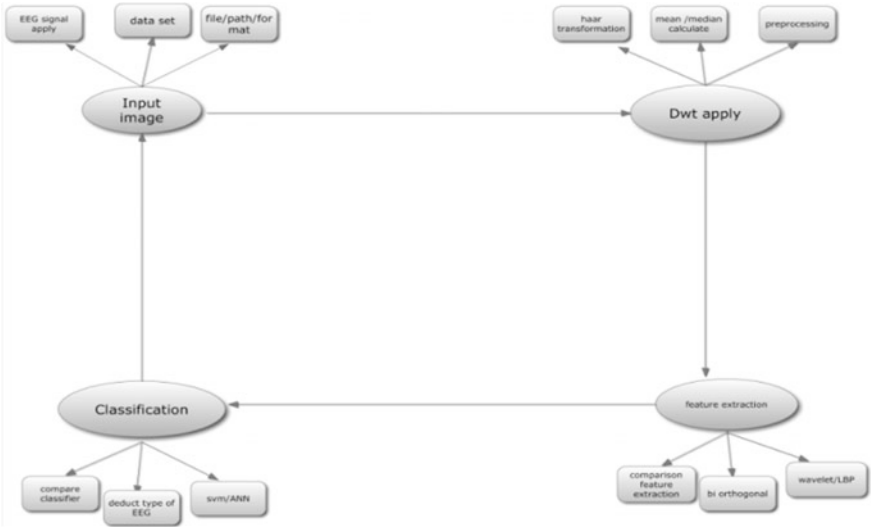


Fig. 2 E–R diagram

Fig. 3 **a** Original image. **b** Grey-scale image. **c** Binary image

(3) To optimize the performance by using a combination of two main drowsiness monitoring methods, i.e., image processing and EEG analysis.

The image of driver can be acquired through image sensors and can be incorporated at appropriate location. Figure 3a shows the input image of the driver, which is split into the red, green and blue planes in order to convert the original image into grey scale image, Fig. 3b, and binary image. Figure 3c shows the binary image is utilized to verify whether the person is sleepy, drowsy or awake.

In order to confirm the sleep related state, we obtain the EEG signal and apply DWT to the signal after filtering. We prefer DWT as it decomposes the signal into sub bands with Multi Resolution Analysis (MRA). The time and frequency domain

Table 1 Maximum values of all the sub-bands of the EEG signal is depicted

Sub-bands	Maximum values (Hz)
γ (gamma)	0.50
β (beta)	0.20
α (alpha)	0.111667
θ (theta)	0.054167
δ (delta)	0.010833

information needs to be obtained simultaneously since EEG is highly non-stationary in time domain and frequency domain as well [8].

We use Daubechies wavelet and five-level decomposition in order to decompose the signal and obtain the detailed and approximate wavelet coefficients. These approximate coefficients possess 97% of energy and are of the maximum value after wavelet decomposition. Further the reconstructed wavelet coefficients are used to compute the frequency analysis of the sub bands using DFT. Table 1 shows the maximum values of sub bands obtained after the frequency analysis of the EEG signal.

Mean, variance, standard deviation and instantaneous power are calculated which will be inputs for classification. LBP gives us the feature extraction values by computing Histogram Oriented Gradients. Haar Wavelet Transform is applied so as compute the feature values of component in the matrix and biorthogonal filter is implemented to get rid of the artifacts. K-means clustering is employed so that 'n' observation can be classified into k classes according to their nearest mean values. This clustering algorithm applies Hebbian algorithm, finding relationship among the input datasets and dividing them into correlated datasets [9]. This research assumes the class of the datasets as unknown. For determining the class these signals belong in, K-means clustering is employed.

$$j = \sum_{j=1}^{k} \sum_{i=1}^{n} \left| |x_i^{(j)} - c_j| \right|^2 \tag{1}$$

where $x_i^{(j)}$ is the feature vector and c_j is the centre of the cluster. This is further processed by SVM, which is a supervised classification algorithm, to classify them into specified classes. This is followed by Pattern recognition Network which finally classifies the clusters into awake, asleep, drowsy, blinking or head-down using a three layered neural net. The first layer has input neurons. The middle layers are hidden layers. Here we give neuron weight as [4 −2] and the bias = −3. The third layer is for the output. The activation function is sigmoid.

$$\varphi(\vartheta) = \frac{1}{1 + e^{-a\vartheta}} \tag{2}$$

By varying a, the slope parameter, the different slopes for the activation function, sigmoid is obtained. It is the most common activation function A, is given by

$$A = p \times w + b \tag{3}$$

If the input matches with the target, it is classified into one of the five classes, awake, drowsy, head down, blinking or asleep. The driver is intimated via a text message displayed in the message box and a voice output to prevent him/her from sleeping. We also intimate the nearest traffic personnel via text message.

5 Results

Figure 4a shows the wavelet reconstruction coefficients of $\gamma, \beta, \alpha, \theta, \delta$. These are the plots of the reconstructed detail coefficients at level 2, 3, 4, 5 and approximate coefficient at level 5, respectively. Figure 4b depicts the plot of the frequency analysis of each of $\gamma, \beta, \alpha, \theta, \delta$. This helps us obtain the peak values of the frequency analysis of each of $\gamma, \beta, \alpha, \theta, \delta$.

The values for column mean, total mean, instantaneous power average power, variance and standard deviation are calculated and tabulated in Fig. 5a. These values act as inputs for the classification stage. Figure 5b is the histograms of oriented gradients obtained after rotation invariant Local Binary Pattern.

Figure 6a depicts the plot for the original EEG signal which is subjected to Haar transformation, the first fluctuation, first trend and the Haar transformed signal. Figure 6b is the stem plot of the maximum values, obtained after subjecting the signal to biorthogonal filtering process.

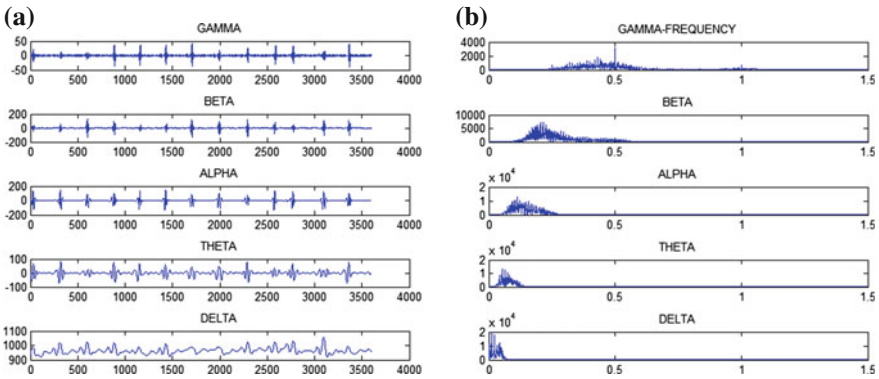


Fig. 4 **a** Reconstructed wavelet coefficients $\alpha, \beta, \gamma, \delta, \theta$. **b** Frequency analysis plots

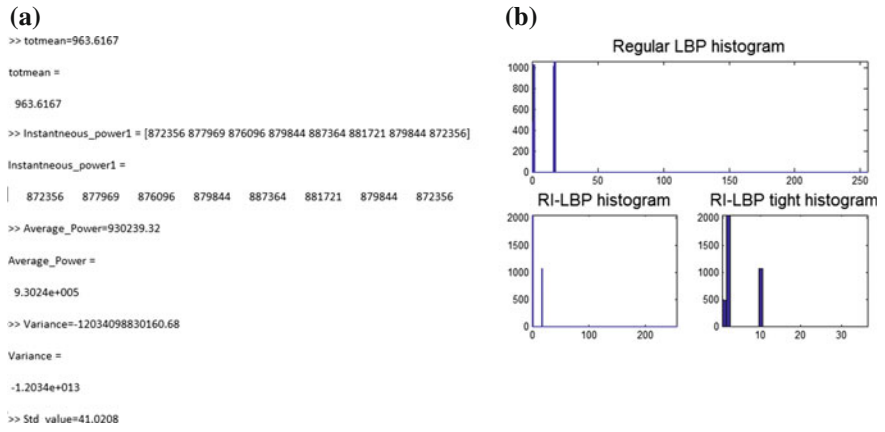


Fig. 5 **a** Calculated values for column mean, total mean, instantaneous power average power, variance and standard deviation. **b** Histogram of oriented gradients

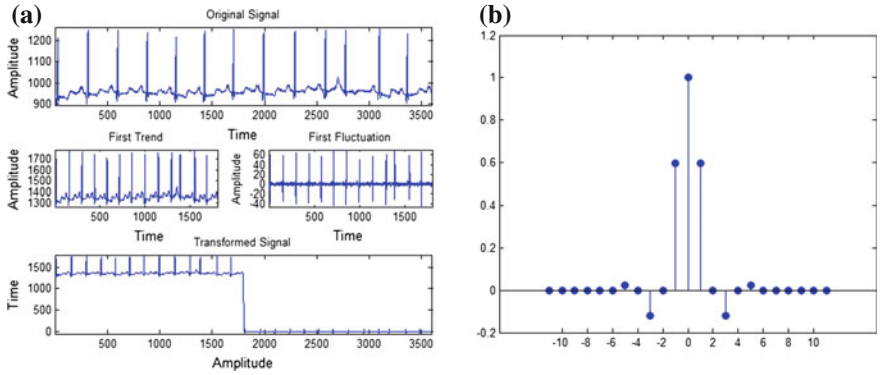


Fig. 6 **a** Plot of the signal, first trend, first fluctuation, transformed signal. **b** Stem plot of maximum values of the filter

The transfer modulus of decomposition low-pass, high-pass filters and reconstruction low-pass, high-pass filters are obtained after biorthogonal filtering process. These have been depicted in Fig. 7a. Figure 7b shows the features classified by k-means clustering being divided into five distinct clusters by the SVM process. These will be input to PRN, which would classify each of these clusters as awake, asleep, head-down, drowsy or blinking state.

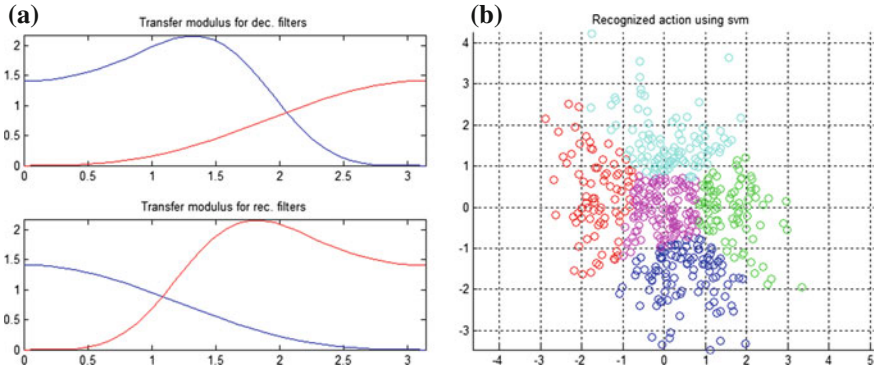


Fig. 7 **a** Transfer modulus for decomposition and reconstruction filters. **b** Recognized action using SVM

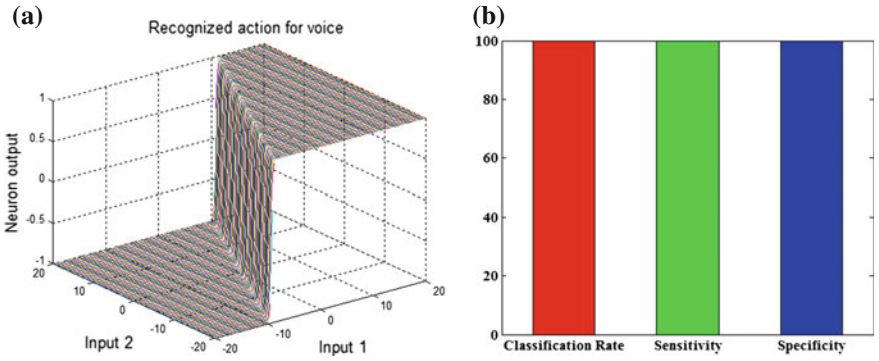


Fig. 8 **a** The activation function sigmoid. **b** Bar graph of performance measures

Pattern recognition Network classifies the clusters into awake, asleep, drowsy, blinking or head-down using a three layered neural net. The first layer has input neurons and the middle layers are hidden layers. Neuron weight is given as [4 −2] and the bias = −3. The third layer is for the output. The activation function is sigmoid. Figure 8a represents the activation function, sigmoid plotted in the mesh grid for the given system. Figure 8b represents the performance measure of the system in the form of a bar graph. Each of classification rate, sensitivity and specificity has been obtained as 100%. This implies that all the datasets that had been used as test features to the system were classified successfully.

6 Algorithm

(1) The driver's image is taken as input and pre-processing and Red, Blue and Green components of the image are obtained. These are used to get grey scale and binary images.
(2) If the person is perceived as sleeping, EEG signal file is loaded and Five-level wavelet decomposition is performed and the detail and approximate coefficients namely cD_2, cD_3, cD_4, cD_5 and cA_5, is obtained.
(3) Wavelet reconstruction is employed for $\gamma, \beta, \alpha, \theta$ and δ based on their frequency content.
(4) Discrete Fourier transform is applied to each of the reconstructed wavelet feature vectors $\gamma, \beta, \alpha, \theta$ and δ for N = 3 and all 5 frequency analyzed results are plotted.
(5) Mean, median, power, variance and standard deviation are evaluated and displayed.
(6) Local Binary Pattern is applied to de-noise the signal, i.e., the driver's EEG. The filter is generated by assigning the filter size = 8 and the radius = 1. Histogram is calculated for the cell for the frequency of all numbers occurring. It can be viewed as a 256-dimensional feature vector. Obsolete values are eliminated, and only relevant ones are mapped.
(7) The Biorthogonal filtering is applied to EEG signal S. Here values for Numerator and Denominator are 3 and 5 which serve the best fit for filtering process. The graphs for decomposition low-pass, high-pass and Reconstruction low-pass, high-pass are plotted.
(8) K-means clustering follows the following algorithm. Dimension = 2. No. of clusters k = 5.

- Start
- x_1, x_2, …, x_n (features) are initialized
- While mean is constant, samples are clustered using estimated means
- For i = 1 to k, distance between each feature is calculated and cluster mean initialized
- If Eq. (1) is satisfied we proceed, else x_i is replaced with all means for cluster i
- End for
- End while
- Stop

(9) Later SVM is applied with the following parameters: Number of train features = 5. Number of labels = 5. Trained features are labeled 1, 2, … n. Test feature is input and compared with each train feature and classified.

(10) Subsequently PRN is applied. Training features and target are set. Input = test feature = 5. Output = 5. Hidden layers = 10. The neural net is created. Division of Data for Training, Validation, Testing is set. The network is trained and tested. Neuron weight = [4 −2]. Bias = −3 for the activation function sigmoid.
(11) MATLAB2013a allows text to speech function for conversion of text intimation into voice intimation.

7 Performance Analysis

If all the test inputs are successfully classified into one of the five classes, then accuracy is said to be 100%. Values of accuracy, sensitivity and specificity of the system are given by:

$$\text{Accuracy} = \text{performance} \times \text{CorrectRate} \times 100$$
$$\text{Sensitivity} = \text{performance} \times \text{Sensitivity} \times 100$$
$$\text{Specificity} = \text{performance} \times \text{Specificity} \times 100$$

Figure 9a shows the tabulated values and Fig. 9b depicts the bar graph of the performance analysis of the system. We compare actual obtained features with the predicted features. An accuracy of 100% implies that all the images that have been used as inputs in the system have been classified into one of the five categories i.e., drowsy, awake, head down, eye blink and asleep.

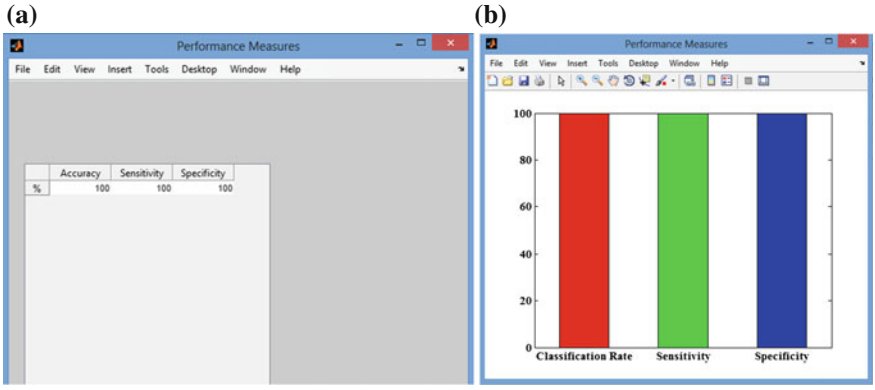


Fig. 9 **a** Tabular column for system performance measures. **b** Bar graph for system performance measures

8 Summary

This research aims to develop an automatic system for classification of EEG signals for sleep detection by analyzing EEG signals of the driver. The wavelet transform is an effective tool to analyze the time as well as frequency components hidden in such non-stationary signals. The first order statistical features are tractable and have low memory requirements. In real world driving scenarios, where manual scoring of EEG signals is impossible and impractical, the K-means clustering technique is a strong decision making tool in order to determine to which sleep related state: (i.e., awake, drowsy, blinking, head down or asleep) each sample belongs. It is concluded that classes obtained by clustering the variance data provides the best separation of clusters when visually inspected. This would be utilized to input the features into a three layer ANN and conduct classification.

References

1. Immrama Institute: The International 10–20 System of Electrode Placement. http://www.immrama.org/eeg/electrode.html. Accessed Feb 2016
2. Physionet: Sleep European Data Format Database. http://www.physionet.edu.in/
3. Peters, B.O., Pfurtscheller, G., Flyvbjerg, H.: Automatic differentiation of multichannel EEG signals. IEEE Trans. Biomed. Eng. **48**(1), 111–116 (2001)
4. Dong, Y., Hu, Z., Uchimura, K., Murayama, N.: Driver inattention monitoring system for intelligent vehicles: a review. IEEE Trans. Intell. Transp. Syst. **12**(2), 596–614 (2011)
5. Polikar, R.: The wavelet tutorial, science magazine's netwatch department. Science **300**(561), 873 (2003)
6. Lin, C.T., Wu, R.-C., Liang, S.-F., Chao, W.-H., Chen, Y.-J., Jung, T.-P.: EEG based drowsiness estimation for safety driving using independent component analysis. IEEE Trans. Circ. Syst. I: Fund. Theory Appl. **52**(12), 2726–2738 (2005)
7. Broughton, R., Hasan, J.: Quantitative topographic electroencephalographic mapping during drowsiness and sleep onset. J. Clin. Neurophysiol. **12**(4), 372–386 (1995)
8. Picot, A., Charbonnier, S., Caplier, A.: On-line automatic detection of driver drowsiness using a single electroencephalographic channel. In: 30th Annual International Conference of the IEEE EMBS (2008)
9. Donalek, C.: Supervised and Unsupervised Learning. http://www.astro.caltech.edu/~donalek/Teaching.html. Accessed 11 Mar 2016

On Blood Flow Through an Overlapping Stenosed Artery

Anuprava Biswas

Abstract In the present study, an analysis of a mathematical model for blood flow through an overlapping stenosed artery is presented by treating the blood as a Newtonian fluid and taking the pressure gradient as a periodic function of time. Perturbation technique is applied to obtain the analytical expressions for velocity profile, volumetric flow rate and wall shear stress by assuming the Womersely parameter to be very small. Effects of overlapping stenosis and hematocrit of red cells on these flow variables are discussed graphically for better understanding of the model.

Keywords Overlapping stenosis · Womersely parameter · Systolic and diastolic pressure

1 Introduction

The study of blood flow through stenotic artery is of great importance as the cause and development of many cardiovascular diseases, particularly atherosclerosis, have been found to be related to the nature of blood movement and mechanical behaviour of vessels walls. Atherosclerosis or stenosis, as it is called, is characterized by an abnormal and unnatural growth that develops in the lumen of the arterial wall at various locations of the cardiovascular system under diseased conditions. With the improvement of the growth, the blood flow in the lumen of the artery is restricted as a result of which individuals may suffer cardiac arrest or stroke.

To understand the effects of stenosis on the arterial blood flow, many investigators (Forrester and Young [1], Halder [2], Misra and Chakravorty [3], Mandal [4], Bali and Awasthi [5] and others) have theoretically studied the pulsatile blood flow through a uniform artery having single mild stenosis. Philip and Chandra [6],

A. Biswas (✉)
Department of Mathematics, Krishnath College Berhampore, Murshidabad, West Bengal, India
e-mail: anuprava0@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_51

Lee and Fung [7] have also investigated the flow of blood through a single stenosed artery and analyzed the effects of slip velocity or peripheral plasma layer thickness on the flow variables such as velocity, wall shear stress, etc. The wall shear distribution plays an important role for describing blood flow characters through arteries. A model of two-layered blood flow through an artery having different shapes of mild stenosis was also considered by Ponalagusamy [8]. However, the stenosis may appear in series form (multiple stenosis) and may be of irregular shapes and so the analysis of the problem of blood flow will be more accurate and interesting in presence of multiple (or overlapping) stenosis. The influence of overlapping stenosis on the arterial blood flow was studied by Chakravorty and Mandal [9], Ismail et al. [10] and Medhavi [11]. Kumar and Aswati [12] investigated the steady and laminar flow of blood with varying viscosity and constant pressure gradient through multiple stenosed arteries. Recently Varshney et al. [13] considered the effect of multiple stenosis on the flow characteristics under the action of transverse magnetic field.

In the present model, an unsteady flow of blood through a multiple stenosed artery having time-dependent pressure gradient and variable viscosity coefficient is considered. Taking the blood to be Newtonian fluid, the effects of hematocrit, multiple stenosis and the amplitude of pressure gradient on the flow characters are obtained and analyzed graphically.

2 Mathematical Formulation

For the development of the mathematical model, let us consider the one-dimensional, axially symmetric, pulsatile, fully developed Newtonian flow through an axisymmetric multiple stenosed arteries. It is also assumed that the blood viscosity varies radially and the density of blood is constant. The basic momentum equations governing the flow in cylindrical co-ordinate system are given by

$$\rho \frac{\partial u}{\partial t} = -\frac{\partial p}{\partial z} - \frac{1}{r}\frac{\partial}{\partial r}(r\tau) \tag{1}$$

$$\frac{\partial p}{\partial r} = 0 \tag{2}$$

Here r and z are the radial and axial co-ordinates respectively, u is the axial velocity, ρ is the blood density and τ is the shear stress defined by

$$\tau = -\mu(r)\frac{\partial u}{\partial r} \tag{3}$$

$\mu(r)$ being the blood viscosity expressed in the form [12]

$$\mu(r) = \frac{\mu_0}{1 - \beta H[1 - (\frac{r}{R_0})^n]} \tag{4}$$

where μ_0 is the coefficient of viscosity of the wall, $\beta = 2.5$, H is the maximum hematocrit, n is the shape parameter and R_0 is the radius of the artery in the non-stenotic region.

The pressure gradient $\frac{\partial p}{\partial z}$ appearing in Eq. (1) is taken as

$$-\frac{\partial p}{\partial z} = A_0 + A_1 \cos \omega t, \; t \geq 0 \tag{5}$$

in which A_0 is the constant amplitude of the pressure gradient, A_1 is the amplitude of the pulsatile component giving rise to systolic and diastolic pressure and $\omega = 2\pi f$, where f is the pulse frequence.

The geometry of the stenosis is described as [11]

$$\frac{R(z)}{R_0} = \begin{cases} 1 - \frac{3}{2}\frac{\delta}{R_0 L_0^4}[11(z-d)L_0^3 - 47(z-d)^2 L_0^2 + 72(z-d)^3 L_0 - 36(z-d)^4], \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad d \leq z \leq d + L_0 \\ 1, \text{ otherwise} \end{cases} \tag{6}$$

where R and R_0 are the radius of the artery with and without stenosis, d indicates the location of the stenosis, L_0 is the length of the stenosis and δ denotes the maximum height of the stenosis appears at $z = d + \frac{L_0}{6}$ and $z = d + \frac{5L_0}{6}$ (Fig. 1).

Now we introduce the following non-dimensional variables for convenience:

$$R' = \frac{R}{R_0}, r' = \frac{r}{R_0}, z' = \frac{z}{R_0}, t' = t\omega, u' = \frac{u}{\omega R_0}, A_0' = \frac{R_0}{\mu\omega} A_0, A_1' = \frac{R_0}{\mu\omega} A_1. \tag{7}$$

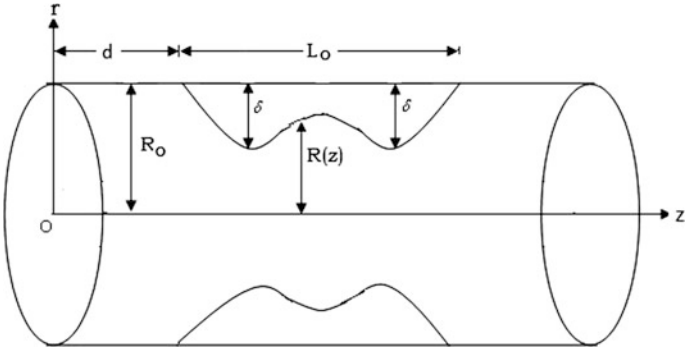


Fig. 1 Geometry of blood vessels with overlapping stenosis

Substituting from (7), Eq. (1) transforms into the dimensionless form as (omitting the primes)

$$\alpha^2 \frac{\partial u}{\partial t} = A_0 + A_1 \cos t + \frac{1}{r}\frac{\partial}{\partial r}\left[r\{1 - \beta H(1 - r^n)\}\frac{\partial u}{\partial r}\right] \tag{8}$$

The boundary conditions in non-dimensional form are

(i) $$u = 0 \text{ at } r = R \tag{9}$$

and

(ii) $$\frac{\partial u}{\partial r} = 0 \text{ at } r = 0 \tag{10}$$

3 Analytical Solutions of the Problem

By considering Womersely parameter to be small ($\alpha^2 < < 1$), we can write the velocity in the following form

$$u(r, z, t) = u_0(r, z, t) + \alpha^2 u_1(r, z, t) + \cdots \tag{11}$$

Substituting (11) into (8), (9) and (10) and then collecting the like terms on both sides, we obtain the following system of equations and the corresponding boundary conditions:

$$0 = A_0 + A_1 \cos t + \frac{1}{r}\frac{\partial}{\partial r}\left(\frac{r}{a_1 + a_2 r^n}\frac{\partial u_0}{\partial r}\right), \tag{12}$$

$$\frac{\partial u_0}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left(\frac{r}{a_1 + a_2 r^n}\frac{\partial u_1}{\partial r}\right), \tag{13}$$

$$u_0 = u_1 = 0 \quad \text{at } r = R, \tag{14}$$

$$\frac{\partial u_0}{\partial r} = \frac{\partial u_1}{\partial r} = 0 \quad \text{at } r = 0. \tag{15}$$

Solving the Eq. (12) with the help of boundary conditions (14) and (15), we get

$$u_0 = \frac{1}{2}(A_0 + A_1 \cos t)\left\{\frac{a_1}{2}(R^2 - r^2) + \frac{a_2}{n+2}(R^{n+2} - r^{n+2})\right\}. \tag{16}$$

Again using the value of u_0, we obtain the solution of (13) by taking into account the boundary conditions (14) and (15) as

$$\begin{aligned} u_1 = \frac{A_1 \sin t}{2}\Big[\frac{a_1^2}{2}\Big\{\frac{3R^4}{16} - \frac{R^2 r^2}{4} + \frac{r^4}{16}\Big\} + a_1 a_2\Big\{\frac{1}{4(n+2)}\big(2R^{n+4} \\ - R^{n+2}r^2 - R^2 r^{n+2}\big) - \frac{n^2+6n+16}{8(n+2)(n+4)^2}\big(R^{n+4} - r^{n+4}\big)\Big\} \\ + \frac{a_2^2}{2(n+2)}\Big\{\frac{(n+3)R^{2n+4}}{(n+4)(n+2)} - \frac{R^{n+2}r^{n+2}}{n+2} + \frac{r^{2n+4}}{(n+4)(n+2)}\Big\}\Big]. \end{aligned} \tag{17}$$

Thus the velocity distribution can be expressed as

$$\begin{aligned} u = \frac{1}{2}(A_0 + A_1 \cos t)\Big\{\frac{a_1}{2}(R^2 - r^2) + \frac{a_2}{n+2}(R^{n+2} - r^{n+2})\Big\} \\ + \frac{\alpha^2 A_1 sint}{2}\Big[\frac{a_1^2}{2}\Big\{\frac{3R^4}{16} - \frac{R^2 r^2}{4} + \frac{r^4}{16}\Big\} + a_1 a_2\Big\{\frac{1}{4(n+2)}\big(2R^{n+4} \\ - R^{n+2}r^2 - R^2 r^{n+2}\big) - \frac{n^2+6n+16}{8(n+2)(n+4)^2}\big(R^{n+4} - r^{n+4}\big)\Big\} \\ + \frac{a_2^2}{2(n+2)}\Big\{\frac{(n+3)R^{2n+4}}{(n+4)(n+2)} - \frac{R^{n+2}r^{n+2}}{n+2} + \frac{r^{2n+4}}{(n+4)(n+2)}\Big\}\Big]. \end{aligned} \tag{18}$$

The volumetric flow rate Q through the artery is given by

$$\begin{aligned} Q &= 2\pi\omega R_0^3 \int_0^R r u dr \\ &= 2\pi\omega R_0^3\Big[\frac{1}{16}(A_0 + A_1 \cos t)\Big\{a_1 R^4 + \frac{4a_2 R^{n+4}}{n+4}\Big\} \\ &\quad + \frac{\alpha^2 A_1 sint}{32}\Big\{\frac{a_1^2 R^6}{3} + a_1 a_2 \frac{2(n+8)R^{n+6}}{(n+4)(n+6)} + \frac{4a_2^2}{(n+3)(n+4)}\Big\}\Big] \end{aligned} \tag{19}$$

Again by using the Eq. (18), the wall shear stress is

$$\begin{aligned} \tau &= \Big[-\mu(r)\omega\frac{\partial u}{\partial r}\Big]_{at\ r=R} \\ &= \frac{1}{2}\mu_0\omega\Big[(A_0 + A_1 \cos t)R + \alpha^2 A_1 \sin t\Big\{\frac{a_1 R^3}{8} + \frac{a_2 R^{n+3}}{2(n+4)}\Big\}\Big]. \end{aligned} \tag{20}$$

4 Numerical Results and Discussions

To obtain a physical insight into the problem, we discuss the above mathematical computations by using the following figures. It is clear from the above analysis that the velocity, flow rate and wall shear stress depend on hematocrit, shape parameter, maximum height of the stenosis, pressure gradient and time respectively. Figures 2, 3 and 4 reveal that the velocity of the blood with radial distance attains the maximum value on the axis and the minimum value at the boundary. It is observed that

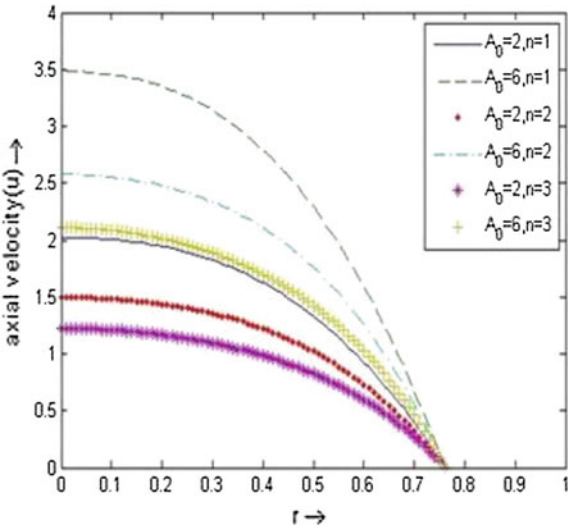


Fig. 2 Variation of axial velocity for d = 0.5, z = 0.75, H = 0.3, $A_1 = 4$, $L_0 = 1$ and t = 0.5

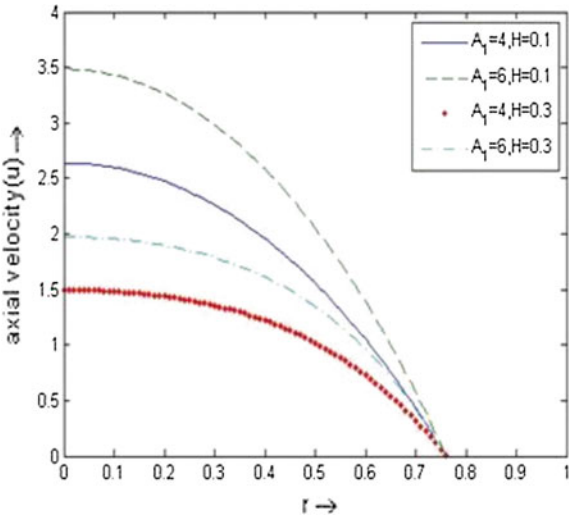


Fig. 3 Variation of axial velocity for d = 0.5, z = 0.75, n = 2, $A_0 = 2$, $L_0 = 1$ and t = 0.5

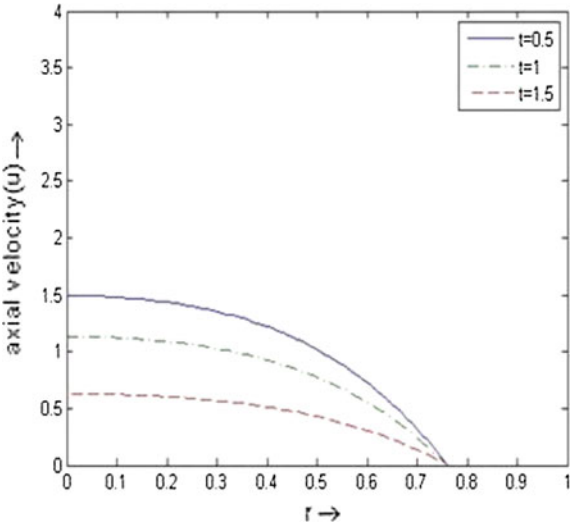


Fig. 4 Variation of axial velocity for d = 0.5, z = 0.75, n = 2, H = 0.3, $A_0 = 2$, $A_1 = 4$ and $L_0 = 1$

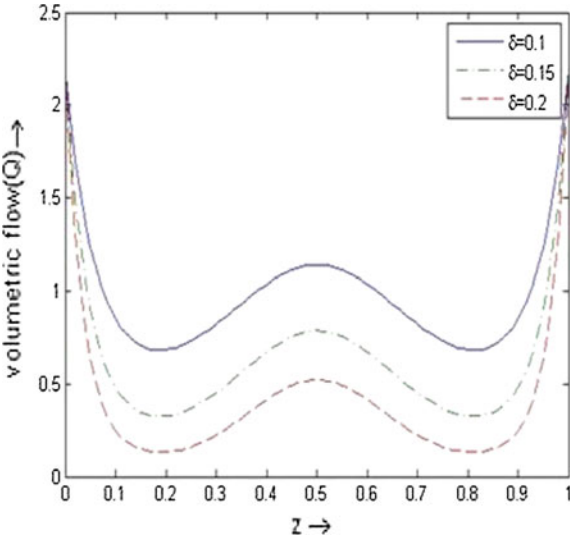


Fig. 5 Variation of volumetric flow rate for d = 0, $L_0 = 1$, n = 1, H = 0.3 and t = 0.5

for a particular value of z and height of stenosis, velocity decreases with the increase of hematocrit, shape parameter and time. Also it is to be noted that as the amplitude of the pressure gradient increases, the velocity increases. Figures 5 and 6 show the variation of the flow rate with axial co-ordinate in the stenotic region for different values of hematocrit, shape parameter and maximum height of stenosis. One observes from Fig. 5 that the volumetric flow decreases with the increase of the maximum height of the stenosis for fixed values of hematocrit, shape parameter, time and length of stenosis. It is also noticed from Fig. 6 that for fixed height and

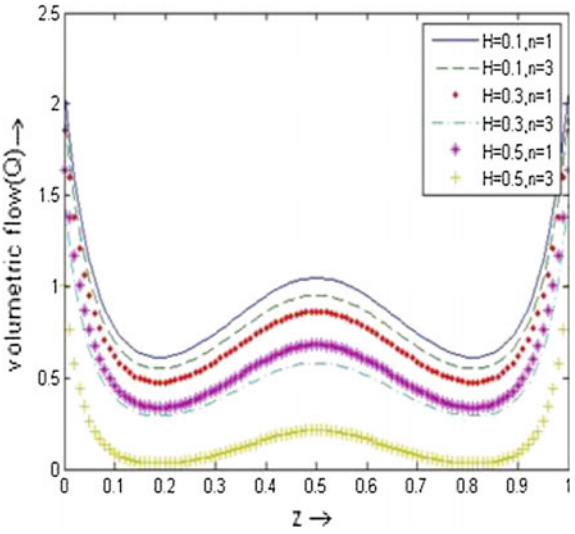


Fig. 6 Variation of volumetric flow rate for d = 0, L_0 = 1, A_0 = 2, A_1 = 4 and t = 0.5

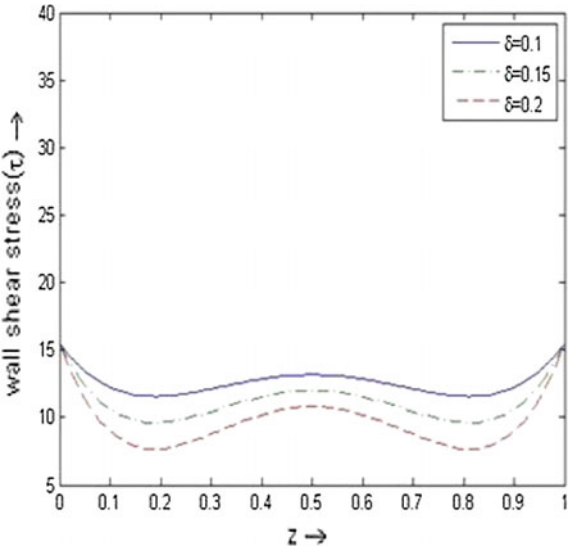


Fig. 7 Variation of wall shear stress for d = 0, L_0 = 1, n = 2, A_0 = 2, A_1 = 4, H = 0.3 and t = 0.5

length of the stenosis, the volumetric flow decreases if hematocrit or shape parameter increases. The variations of wall shear stress with the axial co-ordinate in stenotic region for different values of hematocrit, shape parameter and maximum heights of the stenosis are presented in Figs. 7, 8 and 9. From Fig. 7, it is observed that the wall shear stress decreases with the increasing of the maximum height of the stenosis remaining all the other parameters unaltered. Figures 8 and 9 show that

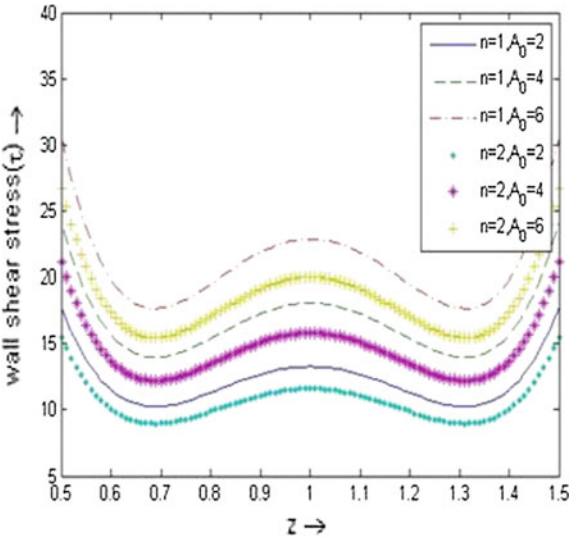


Fig. 8 Variation of wall shear stress for d = 0.5, $L_0 = 1$, $A_1 = 4$, H = 0.3 and t = 0.5

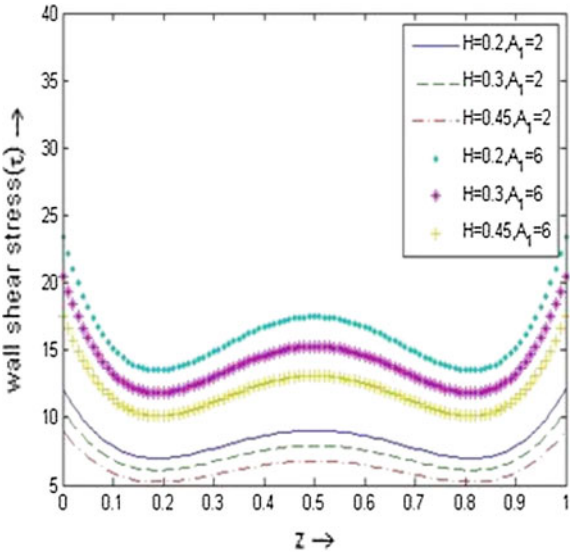


Fig. 9 Variation of wall shear stress for d = 0, $L_0 = 1$, n = 1, $A_0 = 2$ and t = 0.5

for a particular instant the wall shear stress decreases as the value of hematocrit in the red cells and the shape parameter increases but it increases with the increasing of the amplitude of the pressure gradient by keeping the height and length of the stenosis unchanged.

5 Conclusions

The above discussions reveal the significance of present study. The results show that the flow characters vary markedly in the stenotic region. The study also enables one to observe the effects of red cells on the flow characters in presence of overlapping stenosis. Hence, the model will be helpful to us as it provides an insight into physiological situations.

References

1. Forrester, J.H., Young, D.F.: Flow through a converging diverging tube and its implications in occlusive vascular disease. J. Biomech. **3**, 297–316 (1970)
2. Haldar, K.: Effects of the shape of stenosis on the resistance to blood flow through an artery. Bull. Math. Biol. **47**, 545–550 (1985)
3. Misra, J.C., Chakravarty, S.: Flow in arteries in presence of stenosis. J. Biomech. **19**, 907–918 (1986)
4. Mandal, P.K.: An unsteady analysis of non-Newtonian blood flow through tapered arteries with a stenosis. Int. J. Non-Linear Mech. **40**, 151–164 (2005)
5. Bali, R., Awasthi, U.: Effect of magnetic field on the resistance to blood flow through stenotic artery. Appl. Math. Comput. **188**, 1635–1641 (2007)
6. Philip, D., Chandra, P.: Flow of Eringen fluid (simple microfluid) through an artery with stenosis. Int. J. Eng. Sci. **34**, 87–99 (1996)
7. Lee, J.S., Fung, Y.C.: Flow in locally constricted tube at low Reynolds number. J. Appl. Mech. **37**, 9–16 (1970)
8. Ponalagusamy, R.: Blood flow through an artery with mild stenosis: a two-layered model, different shapes of stenoses and slip velocity at the wall. J. Appl. Sci. **7**, 1071–1077 (2007)
9. Chakravarty, S., Mandal, P.K.: Mathematical modeling of blood flow through an overlapping arterial stenosis. Math. Comput. Model. **19**, 59–70 (1994)
10. Ismail, Z., Abdullah, I., Mustapha, N., Amin, N.: A power-law model of blood flow through a tapered overlapping stenosed artery. Appl. Math. Comput. **195**, 669–680 (2007)
11. Medhavi, A.: On macroscopic two-phase arterial blood flow through an overlapping stenosis. J. Sci. Technol. **5**(6), 19–31 (2010)
12. Kumar, A., Awasthi, U.: A mathematical model for blood flow in a multiple stenosis artery. Int. J. Math. Anal. **4**, 2465–2472 (2010)
13. Varshney, G., Katiyar, V.K., Kumar, S.: Effect of magnetic field on the blood flow in artery having multiple stenosis: a numerical study. Int. J. Eng. Sci. Tech. **2**, 67–82(2010)

An Energy-Efficient Congestion Avoidance Priority-Based Routing Algorithm for Body Area Network

Annwesha Banerjee Majumder and Somsubhra Gupta

Abstract In this paper, a routing algorithm for human body area network has been proposed. This algorithm is energy efficient and further, avoids congestions to some extent. One of the main challenges of WBAN is the node life time, i.e. energy level of the nodes. The nodes are selected in this algorithm for communication depending upon three factors that are energy level, number of hops the packet needs to travel while opting that node, and finally the request queue length which is actually the total number of packets traversed through the node for last t time interval. The first factor is for the shortest path to the destination as human body area network consisting of critical data that is needed to be sent fast, however, annexed by other factors as well,namely life time of nodes and congestion-free communication since the shortest path may be congested sometime because of over burden of data. The second factor actually considers the energy level of nodes and the third factor tries to avoid the path which may already been followed by huge numbers of packets.

Keywords Human body area network · Routing algorithm · Energy efficient · Congestion

1 Introduction

The fact is unanimous that continuous health monitoring is essential but these often become infeasible due to scarcity of doctors and hospitals. In extension, the remote location and poor people cannot access round-the-clock health facility directly. Some chronic diseases like cardiovascular diseases, diabetes etc., are gradually emerging as alarming factor in human mortality. Prediction of risks factors of these diseases and continuous monitoring are very much needed. In one of our published

A.B. Majumder (✉) · S. Gupta
JIS College of Engineering, Kalyani, West Bengal, India
e-mail: annwesha.banerjee@gmail.com

S. Gupta
e-mail: gsomsubhra@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_52

papers, a prediction of risk of cardiovascular diseases considering hereditary factors [1] has been proposed. Continuous monitoring of patient of this category is highly required. In the present day context, it is a big challenge to provide health care with limited financial and human resource. Wireless body area network is the answer to this problem. IEEE 802.15.6 is the wireless body area network standard. It is a standard for supporting vast range of data for human body area network. It also provides confidentiality, authentication, integrity, privacy protection and replay defense (Fig. 1).

WBAN is also associated with wireless technologies like ZigBee, WSNs, Bluetooth, cellular networks [2].

Several hospitals are adopting these services. The wireless body area network is now a reality because of advancement of wireless sensor network which is a collection of spatially distributed autonomous devices to monitor different physical and environmental aspects. Body area network is a wireless network of wearable computing devices that can monitor different body responses of human being like EEG, ECG, blood pressure, pulse rate, etc. Routing algorithm plays a very important role in optimum response of BAN.

Along with the human body network, intra-body communication is also a challenging research domain using the redaction device. Red Tacton uses the human body surface as communication path [3]. Human body can be used to send data to several sensors with the help of this technology also.

Several routing algorithm are being proposed in this field. MANET [4] is an energy- efficient routing algorithm for human body area network where it has two phases, in initialization phase all nodes broadcast hello message containing neighbour information and distance to sink node and in the routing phase routes the data with fewer hop count. Dynamic Duty Cycle MAC Algorithm [5] proposed by Jinhyuk Kim et al. is a priority-based energy-efficient routing algorithm that guaranteed low latency. DSDV (Destination Sequence Distance Vector) [6], and reactive protocol DSR (Dynamic Source Routing) [7–10] are two proactive and reactive routing algorithm used in this domain. DSR is based on source routing and DSDV is based on distance vector routing. Okundu et al. has proposed an algorithm containing three processes as link establishment, wakeup service and alarm process.

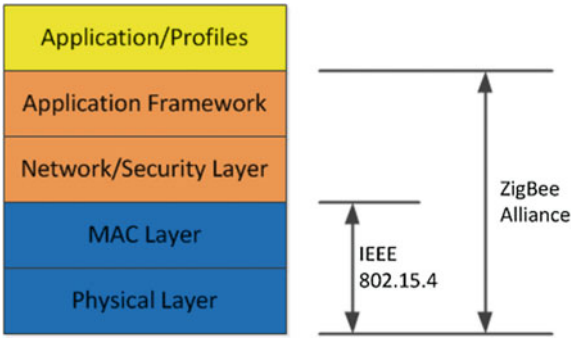


Fig. 1 Layered architecture of wireless body area network

This algorithm is energy efficient and capable of avoiding collisions [11]. TDMA-Based MAC protocol for WBANs called Med MAC [12] proposed by N. F. Timmons et al. consists of two schemes for the power saving: Adaptive Guard Band Algorithm (AGBA) and with Drift Adjustment Factor (DAF). Low Duty Cycle MAC protocol for WBANs [13] can perform analog to digital conversions. Three bandwidth management schemes: Burst, Periodic and Adjust bandwidth is used by B-MAC [14] that reduces the used bandwidth. Time-out MAC (T-MAC) for WBASNs uses flexible duty cycles for increasing energy efficiency [15]. H_MAC is another very interesting algorithm that uses heart beats for synchronization [16]. Reservation-based dynamic TDMA (DTDMA) protocol [17] provides more dependability in terms of lower packet dropping rate and low energy consumption especially for an end device of WBAN. Wise MAC is another MAC algorithm that is scalable and adaptive to traffic load [18]. PACT is an algorithm that is suitable for low delay application [19]. Another dynamic clustering algorithm is LEACH which is a distributed approach [20]. FLAMA is another energy-efficient algorithm that is good for normal traffic [21]. HEED which uses TDMA clustering provides prolonged network lifetime [22].

2 Proposed Method

Through wireless body area networks vast amount of crucial information are transmitted over the networks through many number of nodes. Energy is a very crucial factor for this type of routing algorithm. Very crucial set of data are being transmitted through the nodes that are part of the network. It is a always a better approach not to transmit redundant information again and again, as sending same set of information again and again does not add any effect but and create congestions.

We proposed a routing algorithm for the body area network that is energy efficient. Our algorithm considered three major factors when searching for the path to destination, these are:

i. No. of hops to the destination through the node (HC)
ii. Energy level of the node and (E)
iii. Total no. of packet traversed through the node on last t intervals noted as queue length.

Every node broadcasts their neighbour information along with energy level and queue size in every t time interval as control information along with the data. Every node transmits set of data in next interval only if there are drastically changes in data. This feature will help to reduce congestion over the network. The threshold value is selected depending on the type of information the nodes are sharing.

2.1 Control Information

i. Hop count control information is being shared if there is any changes:

$$HC(N_i, T_i) - HC(N_i, T_{i+1}) \neq 0 \quad (1)$$

ii. Energy level of any node is being shared if there is any changes in the it

$$E(N_i,\ T_i) - E(N_i, T_{i+1}) \neq 0 \quad (2)$$

iii. Queue length information is also being shared based on previous value:

$$QL(N_i,\ T_i) - QL(N_i, T_{i+1}) \neq 0 \quad (3)$$

2.2 Actual Data

Actual data is the most crucial part of the human body area network as based on that data the patients will be taken care.

In actual data field the nodes send human health information: heart beat, blood pressure, pulse rate, body temperature and ECG. All the set of information again need not be sent in every interval as this may increase congestion on the network. If there is any change from the previous interval then only the information is sent again. The nodes for sending actual data are selected sequentially one after another in every time interval, one consideration we have followed is that if any node has the same information as previous interval is will forward to the next node.

$$\text{Select}(N_i, T_i) = \{N_i, T_i | \text{Data}(N_i, T_i) \neq \text{Data}(N_i, T_{i+1}) \quad (4)$$

2.3 Selection of Node for Sending Information Through

We formulate an equation to select the node for transferring data through by considering the weighting factors of the three conditions. We assign weight 0.4, 0.4 and 0.2 for no. of hops, energy level and queue length respectively. This is to be observed that the first condition that is no. of hops to destination through the node is inversely proportional for the selection of node, the second condition that is energy level is also proportional to the selection of the node but the last condition that is queue length is inversely proportional to the selection process as is large numbers of packet traversed through a particular node there is probability of congestion.

Perform the following calculation for choosing nodes that will send data:

$$W1*(1/HC) + W2*E + W3*(1/QL) \quad (5)$$

W1, W2 and W3 are the weight assigned for all above-defined three distinct features those are being considered in our propose method according to their influence in node selection.

And we choose the weight as follows:

$$W1 = 0.4$$
$$W2 = 0.4$$
$$W3 = 0.2$$

1.1 Algorithm:

Our proposed algorithm has following three subparts.

1.1.1 Sharing Control Information

```
IF (HC(Ni, Ti)-HC(Ni,Ti+1) ≠ 0 )
{Share Hop Count information with neighbours}
IF (E(Ni, Ti) - E(Ni,Ti+1) ≠ 0 )
{Share Energy Level information with neighbours}
IF (QL (Ni, Ti) QL(Ni,Ti+1) ≠ 0 )
{Share Queue Length information with neighbours}
```

1.1.2 Selection of Nodes Through Which Data to be Transmitted Depending Upon the Control Information

```
FOR (i = 0 to N)
IF (MAX{(Ni,, Ti), W1*(1/HC) + W2*E + W3*(1/QL)})
{Select Node Ni}
ELSE
{Continue}
```

1.1.3 Selection of Node Sending Actual Data

```
FOR (i = 0 to N)
IF(Data(Ni,Ti) ≠ Data(Ni,Ti+1)
Send Data
Break}
ELSE
Continue}
```

3 Test Result

A sample result of selection of nodes depending on the control information is presented in the following Fig. 2.

The comparison between three factors, viz. energy level, hop count and queue length and thereby estimated value has been presented in the following Fig. 3.

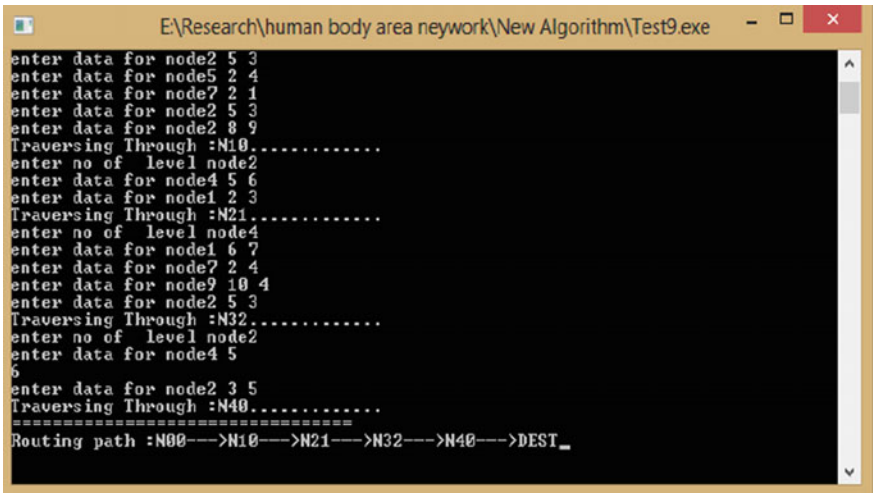


Fig. 2 Test result1

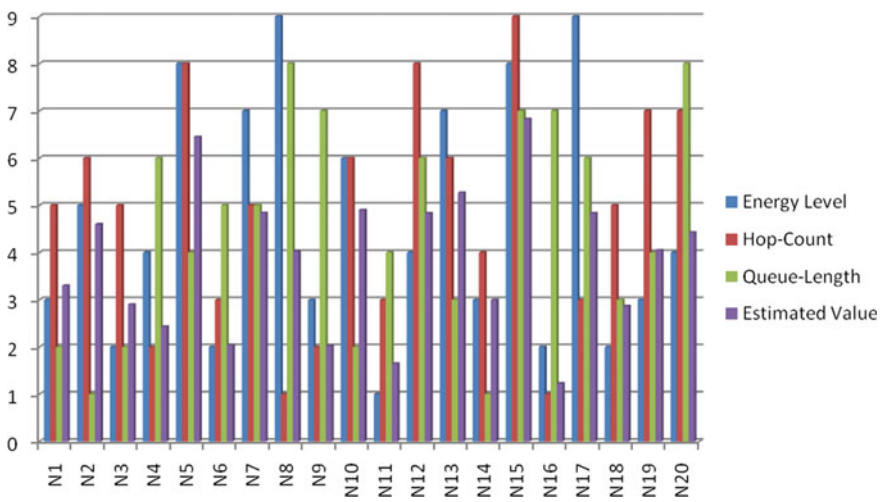


Fig. 3 Comparative selection of nodes depending on three considered values

4 Conclusion

Body area network is an emerging area of research. This has a great contribution to the field of healthcare application and that is possible by the wireless sensors network. As body area network is being used in the remote health monitoring, sending proper data at appropriate time is main focus. And at the same time the life time of the nodes and sensors are also very important for the communication. In the paper, we proposed a method of routing that considers the hop count, energy level and congestion of the nodes also.

So, we can conclude the following from our proposed method:

Node Selection Function:

$$\partial \infty \, 1/\text{Hope Count} \tag{6}$$

$$\partial \infty \, \text{Measure Energy} \tag{7}$$

$$\partial \infty \, 1/\text{Queue Length} \tag{8}$$

References

1. Gupta, S., Banerjee, A.: Proposed intelligent system to identify the level of risk of cardiovascular diseases under the framework of bioinformatics. In: International Conference on Advancements of Medical Electronic. Lecture Notes in Bioengineering, pp. 3–11. doi:10.1007/978-81-322-2256-9_1
2. Kumaria, J.: An energy efficient routing algorithm for wireless body area network. Int. J. Wirel. Microw. Technol. **5**, 56–62 (2015)
3. Gopi, G., Kumar, R.R: Red Tacton—a human area networking technology. Int. J. Eng. Trends Technol. (IJETT) **4**(4), 1310–1315 (2013)
4. Javaid, N., Abbas, Z., Fareed, M.S., Khan, Z.A., Alrajeh, N.: M-ATTEMPT: A new energy-efficient routing protocol for wireless body area sensor networks. In: The 4th International Conference on Ambient Systems, Networks and Technologies (ANT 2013) Procedia Computer Science, vol. 19, pp. 224–231 (2013)
5. Kim, J., Song, I., Jang, E., Choi, S.: A dynamic duty cycle MAC algorithm for wireless body area networks. Int. J. Bio-Sci. Bio-Technol. **4**(2) (2012)
6. The CMU Monarch Project's Wireless and Mobility Extensions to ns, http://www.monarch.cs.cmu.edu/
7. Perkins, C.E., Bhagwat, P.: Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers. In: Proceedings of the SIGCOMM '94 Conference on Communications. Architectures, Protocols and Applications, pp. 234–244, August 1994. A revised version of the paper is available from http://www.cs.umd.edu/projects/mcml/papers/Sigcomm94.ps. Dynamic Source Routing Protocol Internet Draft, http://www.ietf.org/html.charters/manetcharter.html
8. Johnson, D., Maltz, D.: Dynamic source routing in ad hoc wireless networks. In: Imielinski, T., Korth, H. (eds.) Mobile Computing, Chapter 5. Kluwer Academic (1996)
9. Broch, J., Johnson, D., Maltz, D.: The dynamic source routing protocol for mobile ad hoc networks. In: IETF Internet Draft, http://www.ietf.org/internet-drafts/draft-ietf-manetdsr-01.txt (1998)

10. Johnson, D., Maltz, D., Jetcheva, J.: The Dynamic Source Routing Protocol for Mobile Ad hoc networks. In: Internet Draft, https://tools.ietf.org/html/draft-ietf-manet-dsr-07.txt (2002)
11. Omeni, O., Wong, A., Burdett, A.J., Toumazou, C.: Energy efficient medium access protocol for wireless medical body area sensornetworks. In: IEEE (2008)
12. Timmons, N.F., Scanlon, W.G.: An adaptive energy efficient MAC protocol for the medical body area network. In: VITAE (2009)
13. Marinkovic, S.J., Popovici, E.M., Spagnol, C., Faul, S., Marnane, W.P.: Energy-efficient low duty cycle MAC protocol for wireless body area networks. In: IEEE (2009)
14. Fang, G., Dutkiewicz, E.: BodyMAC: energy efficient TDMA-based MAC protocol for wireless body area networks. In: ISCIT (2009)
15. Van Dam, T., Langendoen, K.: An adaptive energy-efficient MAC protocol for wireless sensor networks. In: ACM Conference on Em-bedded Networked Sensor Systems (Sensys), pp. 171–180 (2003)
16. Ullah, S., Shen, B., Riazul Islam, S.M., Khan, P., Saleem, S., Sup Kwak, K.: A study of MAC protocols for WBANs. Sensor (2009)
17. Barati, A., Movaghar, A., Modiri, S., Sabaei, M.: A reliable & energy-efficient scheme for real time wireless sensor networks applications. J. Basic Appl. Sci. Res. **2**(10), 10150–10157 (2012)
18. El-Hoiydi, A., Decotignie, J.D., Hernandez, J.: Low power MAC protocols for infrastructure wireless sensor networks. In: Proceedings of the Fifth European Wireless Conference (EW'04), Barcelona, Spain, pp. 563–569, February 2004
19. Pei, G., Chien, C.: Low power TDMA in large wireless sensor networks. In: IEEE Military Communications Conference (MILCOM), pp. 347–351, October 2001
20. Heinzelman, W.B., Chandrakasan, A.P., Balakrishnan, H.: An application-specific protocol architecture for wireless microsensor networks. IEEE Trans. Wireless Commun. **1**(4), 660–670 (2002)
21. Rajendran, V., Garcia-Luna-Aveces, J.J., Obraczka, K.: Energy-efficient, application-aware medium access for sensor networks. In: Proceedings of 2nd IEEE Conference on Mobile Adhoc and Sensor Systems Conference, Washington, DC, USA, December 2005
22. Younis, O., Fahmy, S.: HEED: A hybrid, energy-efficient, distributed clustering approach for adhoc sensor networks. IEEE Trans. Mob. Comput. **3**(4), 366–379 (2004)

Dynamic Thresholding with Tabu Search for Detection of Hard Exudates in Retinal Image

Diptoneel Kayal and Sreeparna Banerjee

Abstract Diabetic retinopathy is a retinal abnormality found among diabetes patients which may lead to blindness. Hard exudates (HE) is one of the most significant symptoms of DR, which are lipoprotein and cellular debris leaked out of the damaged blood vessels of retina. Detection of HE during early stages helps the patients from loss of vision. This paper proposes a Tabu search-based dynamic thresholding approach to detect HE in retinal images. The proposed method uses Hough transform, median filtering, image thresholding, etc., with a goal to detect HE. The proposed method proves the effectiveness of Tabu search algorithm-based image thresholding technique for the detection of diabetic retinopathy. The proposed method is tested using DIARETDB0 and DIARETDB1 image database and it exhibits a sensitivity of 97.45% and specificity of 96.85%.

Keywords Diabetic retinopathy · Hard exudates · Dynamic thresholding · Tabu search

1 Introduction

Diabetic retinopathy (DR) causes visual impairment among diabetes patients. It has been found that DR is one of the prime causes of blindness among people aged 30–69 years [1]. DR can be classified in two types, namely (1) non-proliferative (NPDR) and (2) proliferative (PDR). NPDR has got various clinical abnormalities, out of which hard exudates are one of the earlier ones. If we could detect hard exudates in earlier stage of the disease, chances of loss of vision could be reduced significantly. Optic disk is a circular bright area of retinal image which could be easily mistaken as exudates, because the brightness is similar to that of exudates.

D. Kayal (✉) · S. Banerjee
Maulana Abul Kalam Azad University of Technology, Kolkata, West Bengal, India
e-mail: diptoneel@gmail.com

S. Banerjee
e-mail: sreeparnab@hotmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_53

So optic disk needs to be detected at very early stages of the method. Later, we have used dynamic thresholding technique in combination with Tabu search algorithm to detect hard exudates. In Sect. 2, we have discussed some earlier methods, in Sect. 3 we have discussed our proposed method, in Sect. 4 we have discussed about the results and testing methodologies adopted. We have calculated the receiver operating characteristic (ROC) curve and confusion matrix, which are given in Sect. 4.

2 Earlier Methods

Various approaches have been proposed to detect hard exudates using image processing techniques. Sagar et al. [2] uses thresholding, edge detection, and histogram specification and contrast enhancement techniques. FCM with colour normalization and contrast enhancement is another image processing method [3]. Morphological operators for hard exudates have been attempted by Osrach et al. [4], Ravishankar et al. [5], Welfer et al. [6] and Kumar et al. [7]. Several authors have additionally used soft computing for classification. These include a multilayer perceptron-based approach by Walter et al. [8], artificial neural network by Garcia et al. [9] and supervised machine learning by Niemeijer et al. [10]. In this paper Tabu search-based dynamic thresholding technique has been discussed to detect hard exudates.

3 Proposed Methodology

RGB image of retinal fundus images has been used as input, and then the image is converted into grayscale. Detection and elimination of optic disk should be done at early stages of the proposed technique. Otherwise it may produce erroneous result. For the detection and masking of optic disk we have used Sobel edge detector and Hough transform. Then median filter is used to eliminate additive white noise of the input image. Now, we have applied image thresholding technique to detect hard exudates. For the detection of the value of thresholding parameter we have used Tabu search algorithm. At the last step, thresholded image and input image needs to be added to get the hard exudates highlighted and to be visible clearly in the output image. Flowchart of the proposed method has been shown in Fig. 1.

3.1 Sobel Edge Detector

Sobel edge detector provides very good performance and it is relatively insensitive to the noise of the image. It can be treated as a 3×3 mask as shown below. Mathematically it uses the equation

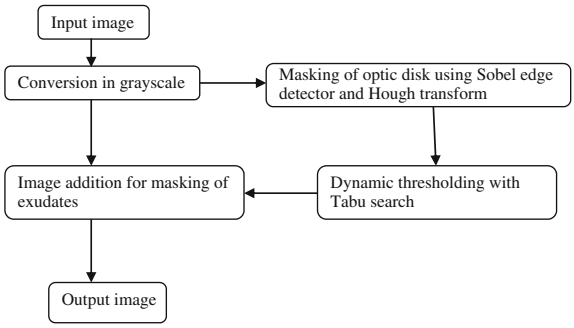


Fig. 1 Flowchart of the proposed method

$$g = \sqrt{G_x^2 + G_y^2}$$

where $G_x = (a_2 + 2a_3 + a_4) - (a_0 + 2a_7 + a_6)$ and $G_y = (a_6 + 2a_5 + a_4) - (a_0 + 2a_1 + a_2)$. In matrix form

$$G_x = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \text{ and } G_y = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

3.2 Hough Transform

Hough transform is very useful for detecting straight lines, curves, and circles since the parametric equation is known to us [11–15]. For detection of circle the equation is as follows,

$$(x - a^2) + (y - b^2) = r^2$$

where a, b are the co-ordinates of the centre of the given circle and r is the radius of the circle [13]. Also,

$$x = a + r\cos\theta \text{ and } y = b + r\sin\theta$$

We have used Sobel edge detector to detect the edges of the input image then applied Hough transform to detect the presence of circular structures in the detected edges. As optic disk is the only circular structure present in the detected edges of the retinal image, we can easily identify and mask it using Hough transform.

3.3 Dynamic Thresholding with Tabu Search

When an image consiss of lighter object on darker background, then we can extract light objects from background using image thresholding operation. Then any point (x, y) in the image at which f (x, y) > T is called an object point, where T is known as threshold parameter. Otherwise the point is called background point. Thresholding operation can be defined as follows [15]:

$$g\ (x, y) = \begin{cases} 1 & if\ (x, y) > T \\ 0 & if\ (x, y) \leq T \end{cases}$$

Success of the thresholding mechanism depends on the value of thresholding parameter (T). Input images may differ in terms of brightness and luminosity. So, if we consider choosing a fixed value for T, it may produce erroneous result. Instead, if we choose to calculate the value of T in runtime, it should result in better segmentation of images. In the proposed method, Tabu search algorithm has been used to find out value of thresholding parameter. In our method we have obtained all the pixel values of an input image and then applied Tabu search [16, 17] onto them to find an optimal value, which in turn acts as T for that input image.

Tabu search (TS) [16, 17] is a neighbourhood search technique to find optimal solution, with the characteristics of local optima avoidance. The basic principle of iteration in TS is to transform solution to a neighbour solution and guide the search into a meaningful direction by utilizing the knowledge of previous iteration. In TS the history (H) of all previously encountered states is maintained. H is also called as Tabu list. During the searching, visit to all the states in H are prohibited in order to prevent re-visiting any previous state. When a neighbourhood N(x) is visited during iteration it is modified to N(H, x). The basic Tabu search algorithm is as follows [16–18],

Initialization
Construct an initial solution S0.
Set S := S0 , f* := f(S0), S* := S0 , H := Ø.
[S = current solution, S* = the best known solution, f* = value of S*]
Search
While termination criterion not satisfied do the following
Select S in argmin [f(S')];
S' $\tilde{N}(S)$
[where N(S) = neighborhood of S, $\tilde{N}(S)$ = non-tabu subset of N(S)]
if f(S) < f*,
then set f* := f(S), S* := S;
Record tabu for the current move in H.
EndOfWhile
Termination criteria

In classical TS the termination criteria could be anything like [16], (1) after a fixed number of iterations, (2) after some number of iterations without an improvement in the objective function value or (3) when the objective reaches a pre-specified threshold value. We have used 500 numbers of iterations as termination criteria.

3.4 Image Addition

Image addition is used to create double exposure of two images into one. If f(m, n) and g(m, n) represents two images then addition of these two images to get the resultant image is given by

$$c(m,n) = f(m,n) + g(m,n)$$

In the proposed method, image addition has been used to highlight and mark the exudates in the input image.

4 Results

The proposed algorithm has been evaluated using 139 retinal fundus images from available diabetic retinopathy image database DIARETDB0 [19] and DIARETDB1 [20]. To evaluate the performance of the proposed method True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) values are considered for calculating sensitivity, specificity and accuracy. Sensitivity, specificity and accuracy are defined as follows.

$$\text{Sensitivity (SE)} = \frac{TP}{TP+FN},\ \text{Specificity} = \frac{TN}{TN+FP} \text{ and Accuracy (AC)} = \frac{TP+TN}{TP+FN+FP+TN}$$

The sensitivity, specificity and accuracy values are found to be 97.45%, 96.85% and 97.05% respectively. Table 1 shows the comparison between various proposed methods in terms of sensitivity, specificity and number of images used. Sample input and output images are shown in Figs. 2a, 3a and 2b, 3b respectively.

ROC curve for the proposed method is shown in Fig. 4. Area under curve is found to be 0.812.

Confusion matrix contains information about actual and predicted classification performed by a classification system. Confusion matrix of the proposed method is given in Table 2.

Table 1 Comparison between various methods

Method	Sensitivity(%)	Specificity (%)	No. of image
Proposed method	97.45	96.85	139
Sagar et al. [2]	99	N.A.	25
Li et al. [3]	100	71	35
Osrach et al. [4]	96	94.6	300
Walter et al. [8]	92.8	92.4	30
Welfer et al. [6]	70.48	98.84	47
Garcia et al. [9]	88.14	92.60	67
Niemeijer et al. [10]	95.00	86.00	300
Ravishankar et al. [5]	95.7	94.2	516
Kumar et al. [7]	97.10	98.30	432

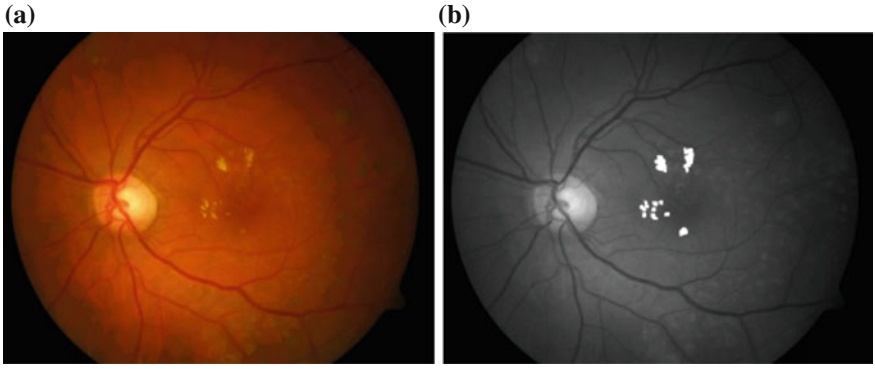


Fig. 2 **a** Input image. **b** Output image

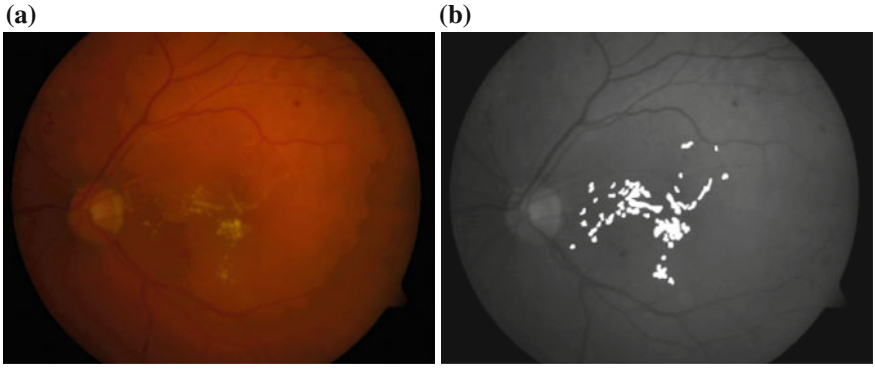


Fig. 3 **a** Input image. **b** Output image

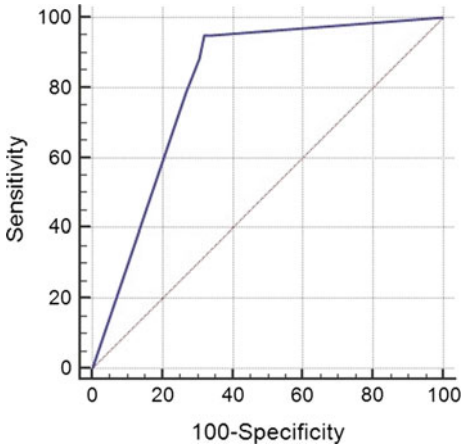


Fig. 4 ROC curve

Table 2 Confusion matrix

		Predicted	
		Negative	Positive
Actual	Negative	54	4
	Positive	2	79

5 Conclusions

The proposed method discussed in this paper exhibits a sensitivity value of 97.25%, specificity value of 96.85% and accuracy of 97.05%. The result has been validated using images from DIARETDB0 and DIARETDB1 standard diabetic retinopathy image database. This result also validates our earlier work using the basic dynamic thresholding algorithm [21].

In future we would like to use more images obtained from medical centres to standardize this algorithm.

Acknowledgements Authors acknowledge a research grant from Department of Biotechnology, Government of India (No. BT/PR4256/BID/7/393/2012) for supporting this research work. They would also want to extend their sincere gratitude to TEQIP-II of Maulana Abul Kalam Azad University of Technology, West Bengal (Formerly known as West Bengal University of Technology) Kolkata for funding of the project.

References

1. Klein, R., Klein, B., Moss, S., Davis, M., Demants, D.: The Wisconsin epidemiology study of diabetic retinopathy type II. Archieve of Opthalmol. **102**(4), 520–526 (1984)
2. Sagar, A.V., Balasubramaniam, B., Chandrasekhara, V.: A novel integrated approach using dynamic thresholding and edge detection for automatic detection of exudates in digital fundus retinal images. In: IEEE International Conference on Computing, pp. 286–292 (2007)

3. Li, H., Chutatape, O.: A model based approach for automated feature extraction in fundus images. In: IEEE International conference on Computer Vision, pp. 127–133 (2003)
4. Osrach, A., Shadgar, B., Markmham, R.: A Computational Intelligence Based Approach for Detection of Exudates in Diabetic Retinopathy. IEEE (2009)
5. Ravishankar, S., Ravishankar, J.A., Mittal A.: Automated feature extraction for early detection of diabetic retinopathy in fundus images. IEEE CVPR'09, pp. 210–217 (2009)
6. Welfer, D., Scharcanski, J., Marinho, D.R.: A coarse-to-fine strategy for automatically detecting exudates in color eye fundus images. Comput. Med. Imaging Graph. **34**(3), 228–235 (2010)
7. Kumar, A.: A segment based technique for detecting exudate from retinal fundus image. Int. J. Comput. Sci. Eng. Technol. **3**(12), ISSN 2229–3345 (2012)
8. Walter, T., Klein, J.C., Massin, P., Erginay, A.: A contribution of image processing to the diagnosis of diabetic retinopathy—detection of exudates in color fundus image of human retina. IEEE Trans. Medical Imag. **21(**10), 256–264 (2002)
9. Garcia, M., Sanchez, C.I., Diez, A., Lopez, M.I., Hornero R.: Detection of hard exudates based on neural networks as a diagnostic aid in the screening for diabetic retinopathy. Telemedicine in Future Health (2006)
10. Niemeijer, M., Ginneken, B.V., Stephen, R., Suttorp-Schulten, M.S.A., Abr`amoff, M.D.: Automated detection and differentiation of drusen, exudates, and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis. IOVS, **48(**5), 2260–2267 (2007)
11. Duda, R.O., Hart, P.E.: Use of hough transform to detect lines and curves in picture. Comm. ACM **15**(1), 11–15 (1982)
12. Ballard, D.H.: Generalizing the hough transform to detect arbitrary shapes. Pattern Recogn. **13**(2), 111–122 (1981)
13. Guil, N., Zapta, E.L.: Lower order statistics and ellipse hough transform. Pattern Recogn. **30** (10), 1729–1744 (1997)
14. Hough, P.V.C.: Methods and means for recognizing complex pattern. US Patent 3069654
15. Loncarie, S.: A survey of shape analysis techniques. Pattern Recogn. **31**(8), 983–1010 (1998)
16. Glover, F.: Future paths for integer programming and links to artificial intelligence. Comput. Oper. Res. **5**, 533–549 (1986)
17. Glover, F., Laguna, M.: Tabu search. In: Reeves, C.R. (ed.) Reeves Modern Heuristic Techniques for Combinatorial Problems, John Wiley & Sons, Inc (1993)
18. Gendreau, M., Potvin J.Y.: Tabu Search. Chapter 6. http://www.inf.ufpr.br/aurora/disciplinas/topicosia2/livros/search/TS.pdf (2016). Accessed 21 Feb 2016
19. DIARETDB0, standard diabetic retinopathy image database. http://www2.it.lut.fi/project/imageret/diaretdb0/ (2016). Accessed 10 March 2016
20. DIARETDB1, standard diabetic retinopathy image database. http://www2.it.lut.fi/project/imageret/diaretdb1/ (2016). Accessed 18 March 2016
21. Kayal, D., Banerjee, S.: An approach to detect hard exudates using normalized cut image segmentation technique in digital retinal fundus image. In: First International Conference on Signal Processing Image Processing and Pattern Recognition, New Delhi, India (2012)

Application of Chaos Game in Tri-Nucleotide Representation for the Comparison of Coding Sequences of β-Globin Gene

Subhram Das, Nobhonil Roy Choudhury, D.N. Tibarewala and D.K. Bhattacharya

Abstract In this paper, we use 2D tri-nucleotide representation based on chaos game theory. We extend the representation from 2D to 3D by taking the third coordinate as the multiple of the first two ones. Complete coding sequences of β globin genes of 10 species are now compared using four types of descriptors—1. Mean of the components of the represented sequences, 2. Standard deviation of the components of the represented sequence, 3. Highest eigen value of M/M matrix and 4. Highest eigen value of J/J matrix. The results in the four cases are critically examined. It is found that the use of J/J matrix with highest eigen value as the descriptor is the best one among the others.

Keywords Chaos game · 2D tri-nucleotide representation · J/J matrix · Highest eigen value

1 Introduction

In bioinformatics, the basic studying strategy for both DNA and protein sequences is to make proper comparisons of both. There are mainly two types of comparison methods—one is based on alignment technique and the other one is based on alignment-free technique. The later one is preferred, as it is less time consuming. Anyway it is mostly based on mononucleotide representation. One such graphical

S. Das (✉) · N.R. Choudhury
Computer Science & Engineering, Narula Institute of Technology, Kolkata, India
e-mail: subhram@gmail.com

N.R. Choudhury
e-mail: nobhonil30390@gmail.com

D.N. Tibarewala · D.K. Bhattacharya
Bio-Science & Engineering, Jadavpur University, Kolkata, India
e-mail: biomed.ju@gmail.com

D.K. Bhattacharya
e-mail: dkb_math@yahoo.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_54

representation is first given by Hamori and Ruskin in 1983 [1]. Graphical representations are found to vary from 2D to 6D. However, directly working with mononucleotides (A, G, C & T) leads to a lot of information loss. So Di- and Tri-nucleotide representations were thought of. The mononucleotide models cannot represent the Di- and Tri- nucleotides without complex calculations [2]. So such representations were found out independently [3–15]. However, the following limitations still remain—1. For 3D representation, the represented values are only 64 in number. Naturally the mean value of such represented coordinates is not of much interest. Even if the cumulative values give much variation, still the use of mean value is not a very satisfactory descriptor. So the final comparison based on means of two types of represented points (normal and cumulative) may not be applicable for comparison of a larger variety of samples [16]. 2. Standard deviation shows the spread of the data rather than determining a theoretical centre, and the cumulative components reduce redundancy. But even the calculations of comparisons based on standard deviation as the descriptor on the cumulative data set is also found to be non-satisfactory [17]. 3. There is always a risk in taking cumulative values, as the resulting time series becomes stochastic. 4. Numerical values and the signs used for tri-nucleotide representation [16, 17] appear to be very much artificial.

In order to avoid these difficulties we have made a very simple approach. We take only the 64 different values obtained by Chaos game representation [18] as the 2D representation of tri-nucleotides and make the representation 3D by taking the third coordinate as the multiple of the first two. As the represented values are now different from those obtained by earlier methods, so to check the improvement in the results, we choose sequentially the descriptors as mean, standard deviation, highest eigen values of M/M matrix and J/J matrix. Compared to the traditional matrix, the J/J matrix can investigate the composition, distribution and chemical properties of bases; it can also picture the biological significance of the sequence [19]. So we try for both M/M and J/J matrix with the expectation that J/J might give better results.

What makes our representation better than the previous ones [1–17] based on tri-nucleotides is that it is a much easier method and hence more efficient. The essential difference lies in getting the 2D tri-nucleotide representations with the help of chaos game theory. The 2D coordinates are obtained in a very natural way using chaos game with only the initial values of the nucleotides. It is known that the exon of β globin genes of different species is essential for pharmaceutical purposes. So we have preferred choosing coding sequences of β globin genes for the purpose of sequence comparison.

2 Methodology

We use the chaos game values shown in the graphical representation of [18] on the non-overlapping triplets of the given DNA sequence for the calculation of different statistical parameters to be used in the analysis of the paper.

Each nucleic acid triplet consists of three coordinates (x, y, z), the first two are obtained from the above chaos game representation, the third one being generated by the multiplication of the first two coordinates. Let N = M/3 be the number of codons in the sequence, where M is the length of the DNA sequence.

Let $x = (x_1, x_2, x_3, \ldots, x_N)$, $y = (y_1, y_2, y_3, \ldots, y_N)$, $z = (z_1, z_2, z_3, \ldots, z_N)$ be the 3D represented points.

Then mean of x, y, z is given by μ_x, μ_y and μ_z respectively, where $\mu_x = \sum_{i=1}^{N} x_i/N$, $\mu_y = \sum_{i=1}^{N} y_i/N$, $\mu_z = \sum_{i=1}^{N} z_i/N$

Then standard deviation of x, y, z is given by V_x, V_y and V_z respectively, where

$$V_x = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \mu_x)^2}{N}},\ V_y = \sqrt{\frac{\sum_{i=1}^{N} (y_i - \mu_y)^2}{N}},\ V_z = \sqrt{\frac{\sum_{i=1}^{N} (z_i - \mu_z)^2}{N}}$$

The M/M matrix is calculated as

$$M_{i,j} = \frac{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}}{|x_i - x_j| + |y_i - y_j| + |z_i - z_j|}$$

The J/J matrix is calculated as

$$J_{i,j} = \frac{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 +}}{|x_i - x_j| + |y_i - y_j|} + \frac{\sqrt{(z_i - z_j)^2}}{|z_i + z_j|}$$

Table 1 The complete coding sequences of β globin genes of 10 species

Species	NCBI ID	Location of each exon	Length of CDS
Human	U01317	19541..19632, 19755..19977, 20833..20961	444
Duck	X15739	291..382, 495..717, 1742..1870	444
Opossum	J03643	467..558, 672..894, 2360..2488	444
Gallus	V00409	465..556, 649..871, 1682..1810	444
Mouse	V00722	275..367, 484..705, 1334..1462	444
Rabbit	V00882	277..368, 495..717, 1291..1419	444
Rat	X06701	310..401, 517..739, 1377..1505	444
Tufted monkey	AY279115	946..1037, 1168..1390, 2218..2346	444
Woolly monkey	AY279114	952..1043, 1174..1396, 2227..2355	444
Hare	Y00347	1485..1576, 1703..1925, 2492..2620	444

Table 2 Distance matrix for the coding sequences of β globin genes of 10 species based on highest eigen value of J/J Matrix

Species	Duck	Gallus	Hare	Human	Mouse	Rabbit	Rat	Tufted monkey	Woolly monkey	Opossum
Duck	0	0.5421	4.532	1.481	2.034	4.781	2.115	3.345	2.768	0.587
Gallus		0	3.990	0.939	1.492	4.239	1.573	2.803	2.226	1.129
Hare			0	3.050	2.498	0.248	2.416	1.187	1.764	5.119
Human				0	0.552	3.299	0.633	1.863	1.286	2.068
Mouse					0	2.746	0.081	1.310	0.733	2.621
Rabbit						0	2.665	1.435	2.012	5.368
Rat							0	1.229	0.652	2.702
Tufted monkey								0	0.577	3.932
Woolly monkey									0	3.355
Opposum										0

We calculate the similarity/dissimilarity between the coding sequences based on the distance matrix measured by

(1) Euclidean distances between three component vectors (μ_x, μ_y, μ_z) of pair of sequences. (2) Euclidean distances between the three component vectors (V_x, V_y, V_z) of pair of sequences. (3) Distance measured by modulus of the difference of highest eigen values of the M/M matrix. (4) Distance measured by modulus of the difference of highest eigen values of the J/J matrix.

The smaller the entry in the distance matrix is, more similar the DNA sequences are. Therefore, we can say that the distances between evolutionary closely related species are smaller, while those between evolutionary distant species are larger. We draw the phylogenetic tree based on similarity/dissimilarity matrix using UPGMA in MEGA4 software [20].

3 Result and Discussion

Table 1 shows the information regarding corresponding sequences of 10 different species and Table 2 shows the distance matrix of the complete coding sequences of β globin genes of 10 different species based on highest eigen value of J/J matrix. Distance matrix using Euclidian distance and corresponding Phylogenic trees are also obtained similarly in other three cases. We observe that the phylogenetic tree Fig. 1 using their highest eigen value of J/J matrix generates the best result among others. From Fig. 1 we also observe that the more similar species pairs are like Mouse—Rat, Tufted Monkey—Woolly Monkey, Hare—Rabbit, Gallus—Duck are come closer to each others. Our phylogenetic tree agrees with that found in [16] for the species taken in common.

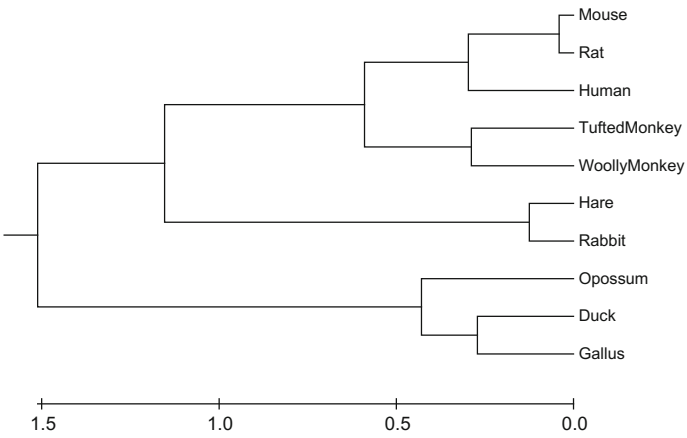


Fig. 1 Phylogenetic tree of 10 different species based on their complete coding sequence of β globin genes using their highest eigen value of J/J matrix

4 Conclusion

In this paper, we propose a new method of nucleotide representation using chaos game theory. By comparing four descriptor, we conclude that highest eigen value for J/J matrix is the best among the four descriptors: mean, standard deviation, highest eigen value for M/M matrix and highest eigen value for J/J matrix for comparison of complete coding sequences of β globin genes for the above 10 species. We therefore conclude that our method is effective for evaluating sequence similarities on an intuitive basis. However, our method is experimented on only 10 different sequences; in near future we like to apply our method on large numbers of species.

References

1. Hamori, E., Ruskin, J.: H curves, a novel method of representation of nucleotide series especially suited for long DNA sequences. J. Biol. Chem. **258**, 1318–1327 (1983)
2. Guo, F.B., Ou, H.Y., Zhang, C.T.: ZCURVE: a new system for recognizing protein–coding genes in bacterial and archaeal genomes. Nucl. Acids. Res. **31**, 1780–1789 (2003)
3. Zhang, C.T., Zhang, R.: Analysis of distribution of bases in the coding sequences by a diagrammatic technique. Nucl. Acids Res. **19**, 6313–6317 (1991)
4. Zhang, R., Zhang, C.T.: Z curves, an intuitive tool for visualizing and analyzing the DNA sequences. J. Biomol. Struct. Dyn. **11**, 767–782 (1994)
5. Nandy, A.: A new graphical representation and analysis of DNA sequence structure: I. Methodology and application to globin genes. Curr. Sci. **66**, 309–314 (1994)
6. Randic, M., Vracko, M., Lers, N., Plavsic, D.: Novel 2–D graphical representation of DNA sequences and their numerical characterization. Chem. Phys. Lett. **368**, 1–6 (2003)
7. Randic, M., Vracko, M., Zupan, J., Novic, M.: Compact 2–D graphical representation of DNA. Chem. Phys. Lett. **373**, 558–562 (2003)
8. Liao, B., Wang, T.M.: Analysis of similarity/dissimilarity of DNA sequences based on 3–D graphical representation. Chem. Phys. Lett. **388**, 195–200 (2004)
9. Randic, M.: Graphical representations of DNA as 2–D map. Chem. Phys. Lett. **386**, 468–471 (2004)
10. Liao, B., Wang, T.M.: 3–D graphical representation of DNA sequences and their numerical characterization. J. Mol. Struct. (Theochem) **681**, 209–212 (2004)
11. Chi, R., Ding, K.Q.: Novel 4D numerical representation of DNA sequences. Chem. Phys. Lett. **407**, 63–67 (2005)
12. Yao, Y.H., Nan, X.Y., Wang, T.M.: A new 2D graphical representation—Classification curve and the analysis of similarity/dissimilarity of DNA sequences. J. Mol. Struct. (Theochem) **764**, 101–108 (2006)
13. Liao, B., Ding, K.Q.: A 3D graphical representation of DNA sequences and its application. Theor. Comput. Sci. **358**, 56–64 (2006)
14. Song, J., Tang, H.W.: A new 2–D graphical representation of DNA sequences and their numerical characterization. J. Biochem. Biophys. Methods **63**, 228–239 (2005)
15. Zhang, Z.J.: DV–Curve: a novel intuitive tool for visualizing and analyzing DNA sequences. Bioinformatics, vol. 25, pp. 1112–1117 (2009)
16. Yu, J., Wang, J., Sun, X.: Analysis of similarities/dissimilarities of DNA sequences based on a novel graphical representation. MATCH Commun. Math. Comput. Chem. **63**, 493-512 (2010)

17. Das, S., Palit, S., Mahalanabish, A.R., Choudhury, N.R.: A new way to find similarity/dissimilarity of DNA sequences on the basis of dinucleotides representation. In: Computational Advancement in Communication Circuits and System, pp. 151–160. Springer (2015)
18. Randic, M., Zupan, J., Balaban, A.T.: Unique graphical representation of protein sequences based on nucleotide triplet codons. Chem. Phys. Lett. **397**, 247–252 (2004)
19. Luo, J., Guo, J., Li, Y.: A new graphical representation and its application in similarity/dissimilarity analysis of DNA sequences. In: 4th International Conference on Bioinformatics and Biomedical Engineering (2010). doi:10.1109/ICBBE.2010.5515203
20. Kumar, S., Nei, M., Dudley, J., Tamura, K.: MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief. Bioinform. **9**, 299–306 (2008)

Computational Methodologies Followed in Structure Based In-Silico Drug Design: An Example

Indrani Sarkar and Sanjay Goswami

Abstract With the emergence of multidrug resistant mycobacterium tuberculosis, it is an immediate necessity to design new potent drugs. One step toward it is to understand the mechanism of inhibition of the enzyme InhA (trans-2-enoyl-ACP-reductase), which has an important function in the building of micobacterial cell wall. An attempt through molecular modeling and simulation methods is presented here to explain binding affinities of different derivatives of drugs and also the role of some important amino acid residues in the active site of the InhA.

Keywords Molecular modeling · Molecular dynamics · M. tuberculosis · Enoyl acyl carrier protein reductase · Carboxamides · Binding free energy

1 Introduction

Despite more than five decades of effective chemotherapy tuberculosis caused by Mycobacterium tuberculosis (MTB) still remains an important problem of public health. Infection with drug-sensitive strains of MTB can be effectively treated with a combination of drugs known as multi-drug therapy. The drugs used are isoniazid, rifampicin, pyrazinamide, and ethambutol or streptomycin. According to WHO 79% of multi-drug resistance (MDR)-TB cases are now due to super strains. They are resistant to at least three of the main four drugs used to treat tuberculosis. It is, therefore, urgent to identify the mechanism of drug action and resistance. This

I. Sarkar (✉)
Department of Physics, Narula Institute of Technology, 81, Nilgunj Road, Agarpara, Kolkata 700109, West Bengal, India
e-mail: indrani.sarkar.dumdum@gmail.com

S. Goswami
Department of Computer Applications, Narula Institute of Technology, 81, Nilgunj Road, Agarpara, Kolkata 700109, West Bengal, India
e-mail: sanjaygoswamee@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_55

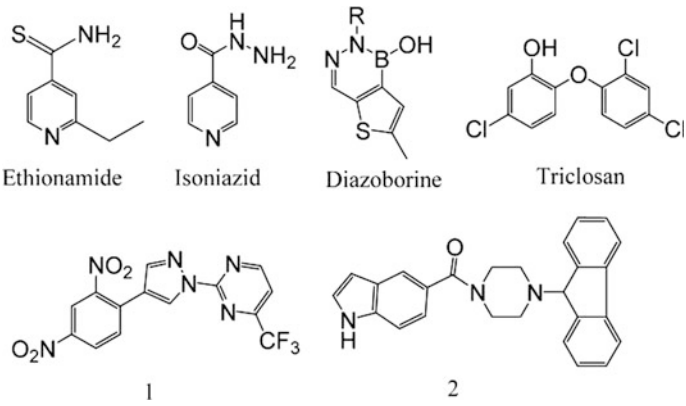


Fig. 1 Inhibitors of InhA

understanding should contribute to the design of more potent antimicobacterial agents for controlling drug sensitive and drug-resistant TB. The enzymes which are involved in the fatty acid synthetic pathway are potent targets for the design of new antibacterial agents. Mycobacterium tuberculosis contains unique fatty acids known as the micolic acids. They constitute the micobacterial cell wall. The InhA enzyme or trans-2-enoyl—ACP-reductase (EC1.3.1.9) carries out the synthesis of micolic acids which are unusually long chain alpha-alkyl, beta-hydroxy fatty acids of 60–90 carbons. The TB- specific drugs isoniazid and ethionamide (Fig. 1) inhibit the enzyme and stop the formation of cell wall by inhibiting mycolic acid synthesis. Isoniazid does not inhibit the enzyme directly. It is converted to an activated form of the drug by the catalase peroxidase (Kat G) of the micobacterium. The activated form of the drug then attach itself covalently to the nicotinamide ring of the NADH bound within the active site of the enzyme (Fig. 2) [1–3]. Mutations in the KatG enzyme prevents the process of conversion of INH into the activated form, hence the drug becomes ineffective. So compounds which directly inhibit the InhA enzyme without requiring activation by KatG are considered as very promising new drugs against tuberculosis. Many new compounds functioning as direct inhibitors are being discovered [4–7]. Consistent with these results X-ray crystallographic studies have revealed that all mutations in the enoyl reductase causing INH resistance were located in the nucleotide-binding site. The findings suggested that the resistance toward INH is due to the specific interaction between enzyme and NADH cofactor within the enzyme-binding site.

1.1 Identification of Novel Inhibitors

To understand the mechanism of InhA full characterization of the binding pocket of the enzyme is necessary which has been done by X-ray crystal structure analysis.

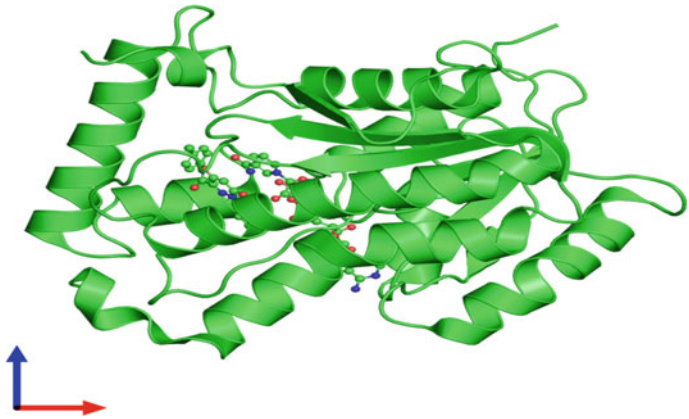


Fig. 2 The mode of interaction of the inhibitor inside the active site of the M. Tuberculosis InhA enzyme is shown in secondary structure and the inhibitor is shown in ball–stick model (Entry 4dor, protein data bank in Europe)

The crystal structure of InhA complexed with NADH has been solved and deposited in the Protein Data Bank (PDB). The binding site can accommodate long-chain substrates which support the broad substrate specificity of InhA. A new class of direct inhibitors, pyrrolidine carboxamide shows good inhibitory action toward InhA. Other known drugs are diazoborines and triclosan but both of these are nonselective and relatively weak agents. Two series of InhA direct inhibitors are pyrazole derivative (1) and indole-5-amides (2) (Fig. 1). They have shown both in vivo and in vitro activity. The pyrrolidine carboxamide series was chosen for modeling studies due to their availability and easy chemical synthesis.

2 Methodology

2.1 Modeling of InhA-Inhibitor Complexes by Docking

The co-ordinates for the crystal structure of the enzyme (InhA) were taken from Protein Data Bank in Europe (Entry 4dor) with highest crystallographic resolution (Fig. 2). The model was validated by correcting any improper bond lengths, bond angles, torsion angles, missing atoms, and connectivity within the atoms. The inhibitors with their inhibitory activities were collected from the literature. All the inhibitors were energy minimized by using CHARMM forcefield as available in Discovery Studio suite of software. The refined structure of InhA was used for exploring the potential active sites. The inhibitors were docked into the active site of the InhA with the LibDock program. The conformation of the inhibitor poses were generated using BEST algorithm and then placed into binding site spheres based on triplets matching hotspots. The 30 top-ranked score poses were selected

for calculation of binding energy between the enzyme and an inhibitor. Energy minimization was performed to remove any enzyme-inhibitor van der waals clashes prior to calculating the binding energy. 500 steps of Steepest Descent with free movement of atoms within the binding site sphere were used during minimization. The receptor inhibitor complex pose with the lowest binding energy was selected for further calculation and analysis.

2.2 Molecular Dynamics Simulation

Molecular dynamics simulations were performed with the CHARMM force field. The enzyme-inhibitor complexes were solvated with TIP3P water molecules in an explicit spherical boundary condition. Each simulation was carried out in two steps. At first energy minimization was carried out by Steepest Descent and Conjugate Gradient methods to remove short contacts. Subsequent heating, equilibrium, and production run was carried out by molecular dynamics simulation for 10 ns at 300 K. The time step chosen for production run was 2 fs. The trajectories were saved every 20 ps interval for analysis.

3 Results and Discussions

To identify lead compounds Cambridge Database for small molecules was searched. Some 10 inhibitors were identified. The inhibitor stabilizes by forming a hydrogen-bonding network with the enzyme active site residues and the cofactor. The hydrogen-bonding network is also observed in the crystal structures of inhibitors complexed with InhA. The substituents on the phenyl ring and the cyclohexyl ring while keeping the core features of the molecule have effect on its activity [8–11]. The modeled structures of the inhibitors were examined and following features were found (Table 1).

It was seen that the hydrogen-bonding pattern of the drug with the cofactor and side-chain of Tyr158 is conserved for all the inhibitors.

Table 1 Inhibitors with their calculated binding energies

Molecular formula	Molecular weight	Number of hydrogen donors	Number of hydrogen acceptors	Binding energy Δkcal/mol
$C_{19}H_{16}N_4O_2S$	364.43	1	6	−9.20
$C_{19}H_{15}ClN_4O_2S$	398.87	1	6	−8.76
$C_{20}H_{18}N_4O_2S$	378.45	1	6	−7.45
$C_{20}H_{18}N_4O_3S$	394.45	1	7	−9.60
$C_{19}H_{15}N_5O_4S$	409.42	1	9	−9.86

The mutation in the Gly-rich loop results in the decrease or elimination of enzymic activity, the resistance of the enzyme, and decrease of cofactor affinity of wild-type inhA. Two mutations in the Gly-rich loop (I21V and I16T) were examined and the molecular aspects of NADH affinity differences were identified.

Tyrosine 158 and Lysine 165 are two important residues involved in the catalytic mechanism of InhA. Their role has been studied experimentally. Replacement of Tyr 158 with phenylalanine (Y158F) and alanine Y158A decreases the binding constant.

Experimental observation was that when Lys165 was replaced with glutamine (K165Q) and arginine (K165R) it had no effect on the ability of InhA to bind with NADH. Lysine 165 has involvement in cofactor binding [9].

The interaction energy was calculated between inhibitors and particular amino acids within a few of each inhibitor. This was done to show the contribution of each crucial residue for the binding affinities of the inhibitor. Binding-free energies for the caboxamide complexes were computed by LIE method to gain quantitative idea about the binding affinities of the inhibitors. The MD simulation findings were at par with the experimental results. The structural requirement of the carboxamides to bind to InhA with improved selectivity was attempted in this work. The difference between the binding mode of dichloro compounds and the singly substituted 3-Br compounds was also observed [10, 11].

MD simulation for the carboxamide-complex with inhA will give an insight into the dynamic behavior of the interaction of drugs with the active site of InhA.

Acknowledgements The author is thankful to the University Grant Commission, Government of India for providing necessary facilities.

References

1. Rozwarski, D.A.: Crystal structure of the Micobacterium tuberculosis enoyl-ACP reductase, InhA in complex with NAD+ and a C16 fatty acyl substrate. J. Biol. Chem. **274**, 15582–155589 (1999)
2. Luckner, S.R.: X ray crystal structure of a slow, tight binding inhibitor of InhA from M. Tb. J. Biol. Chem. **285**, 14330 (2010)
3. He, X., et al.: Pyrrolidine carboxamides as a novel class of inhibitors of InhA from M.Tb. J. Med. Chem. **49**(21), 6308–6323 (2006)
4. Punkvang, A., et al.: Investigating the structural basis of arylamides to improve potency against M. Tuberculosis strain through molecular dynamics simulation. Eur. J. Med. Chem. **45**, 93–5585 (2010)
5. Schroeder et al.: Molecular dynamics simulation studies of the wild type I 21 V and I 16T mutants of isoniazid resistant M. Tb InhA in complex with NADH, Towards understanding of NADH–InhA different affinities. Biophys. J. **89**, 876–884 (2005)
6. He, X., et al.: Inhibition of the MTb enoyl acyl carrier protein reductase InhA by arylamides. Bioorg. Med. Chem. **15**(21), 6649–6658 (2007)
7. Subba Rao, G., Vijayakrishnan, R., Kumar, M.: Structure based design of a novel class of potent inhibitors of InhA, the enoyl acyl carrier protein reductase from micobacterium tuberculosis: a computer modeling approach. Chem. Biol. Drug Des. **72**, 444–449 (2008)

8. Kumar, M., Vijayakrishnan, R., Subba Rao, G.: In silico structure based design of a novel class of potent and selective small peptide inhibitor of M.Tb. Dihydrate reductase, a potential target for anti T.B. drug discovery. Mol. Divers. **14**(3), 594–604 (2010)
9. Muralidharan, J., Suguna, K., Surolia, N.: Exploring the interaction energies for the binding of hydroxydiphenyl ethers to enoyl acyl carrier protein reductases. J. Biomol. Struct. Dyn. **20**, 589–94 (2003)
10. Kini, S.G., Bhat, A.K.: Synthesis, antitubercular activity and docking study of novel cyclic azole substituted diphenyl ether derivatives. Eur. J. Med. Chem. **44**, 492–500 (2009)
11. Kouassi, A.F., et al.: Computer-aided design of orally bioavailable pyrrolidine carboxamide inhibitors of enoyl-acyl carrier protein reductase of *Mycobacterium tuberculosis* with favorable pharmacokinetic profiles. Int J. Mol. Sci. **16**(12), 29744–29771 (2015)

A Study on Data Acquisition and Analysis in Real Time for Prediction of Cardiovascular Disease in Real Time Using Indigenous Low-Cost CHMS

Meghamala Dutta, Deepneha Dutta, Swati Sikdar, Sourav Dutta and Shreya Dutta

Abstract In a developing country like India with a majority of the population living in the rural areas sudden and fatal heart attacks are very common amongst the age group of 30–60 years. Studies have established that subdued cardiovascular abnormalities often go unnoticed in the ECG in the early days and clinical intervention is sought only during emergency conditions. Till date, the recordings of ECG along with heart rate, heart sounds and PPG are still the 'only' confirmatory diagnostic tools for cardiovascular abnormalities. This problem increases many folds especially in the rural areas where lack of awareness towards health issues gets multiplied with cost factors and lack of adequate medical facilities for the populace. Prior to design and implementation of CHMS our work was committed to the hardware unit CHMS (Cardiac Health monitoring System in Real time) but further advancements and coupling with the handheld user-friendly mobile phone has prompted us name our indigenously designed Mobile ECG Acquisition Device as CHMS. CHMS is a low-cost predictive tool which enables the subject to record and store electrical signals from the heart in a simple user-friendly and readable format for future reference. In this paper, we have analyzed the data collected by the CHMS in real time from consenting adult volunteers. The objective was to establish the accuracy of data acquired by the CHMS. A comparative analysis with the "standard" commercial Recorder has also been represented. Standard Deviation has been calculated to be within limits.

Keywords CHMS · CVD · ECG · Electrical signals · Heart sound · LPF · PPG

M. Dutta (✉) · D. Dutta · S. Sikdar
Department of Biomedical Engineering, JIS College of Engineering, Kalyani, India
e-mail: megh0333@hmail.com

S. Dutta
IBM India Ltd., New Delhi, India

S. Dutta
IIMSAR & Dr. BCRHH, Haldia, India

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_56

1 Introduction

WHO statistics (2005–2012) on cardiovascular disease (CVD) reveals that subdued and chronic CVD kills 1 person every 5 s and each year 7.6 millions of death occurs only due to the subdued CVD. The following distribution shows the percentage of people suffering from different types of heart diseases [1] (Fig. 1).

A thorough understanding of the nature of various cardiovascular disease reveals the association of physiological abnormalities and subdued physiological disorders in the human beings that might have risen due to stress, lifestyle related disorders, cigarette smoking, physical inactivity, increased body mass index (BMI), pathological conditions etc. These are the key risk factors of premature cardiovascular disease (CVD) and possible death especially in the age group 35–75 years [2]. Related disorders, cigarette smoking, physical inactivity, increased body mass index (BMI), pathological conditions etc. These are the key risk factors of premature cardiovascular disease (CVD) and possible death especially in the age group 35–75 years [2].

Infrastructure for carrying out the tests thus making With the popularity and essentiality of health care, continuous advancements in medical technology have played a key role in coming up with user-friendly devices and procedures. The daily monitoring of health-status-related parameters is getting more and more valued. Early detection is possible in developed countries with expensive and advanced machineries, which can keep the data recorded in a user-friendly device affordable by the patient. The existing technology for these purpose are high-end like those that are coupled with customized expensive attachments and even more cumbersome say for example, Holter Monitor etc. This problem increases manifold especially in developing countries like ours, where lack of awareness towards health issues gets multiplied with cost factors and lack of medical facilities. Till date, the recordings of ECG along with heart rate, heart sounds and PPG are still the 'only' confirmatory diagnostic tools for cardiovascular abnormalities. However, timely monitoring and analysis can predict subdued cardiovascular disorders [3–5]. But ECG is clinically effective when the disturbances of the cardiovascular system are reflected in the recording. But often at an early stage of CVD, recordings in an

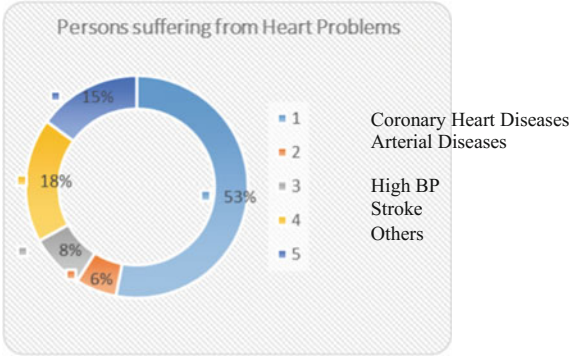


Fig. 1 WHO studies 2005–12

ECG are near normal and the physician fails to diagnose and treat CVD. Hence, other physiological signals like PPG (for measuring heart rate), heart sounds should be seriously considered to be an important clinical marker for diagnosis and treatment of CVD in addition to ECG [6, 7].

The widely used ECG studies is limited for detection of abnormalities rising due to disturbances in the conduction of electrical impulse across the heart musculature while heart sound studies, e.g. the phonocardiograph or echocardiograph deal with the mechanical defects of the heart. Even the photoplethysmograph recording can also give us the preliminary indications for the sudden abnormalities [8]. However, these diagnostic tools have their own limitations, as they require skilled medical personnel for interpretation of result, in laboratory or hospital the tests too expensive and are condition specific. Hence, cases of mild or subdued cardiac abnormalities often go unnoticed and results in undertreatment. Development of a low-cost simple user-friendly device for continuous monitoring of physiological parameters is an urgent need especially for the rural Indian population. To satisfy these different needs development of a comfortable and robust procedure involving device design, interfacing with the handheld mobile device, programming for storage and interpretation of acquired signals is therefore an important area of concern. Phonocardiogram (PCG) is a physiological signal reflecting the cardiovascular status. It could be recorded by microphone equipped in-hand devices, like the smartphone, even without direct skin contact. However, high inter and intra-variance of PCG make its processing challenging. For PCG-based HR measurement, a robust method is still strongly required. That is why in this work we have dealt with mainly ECG signal processing and recording along with PPG signal processing and recording. The recording of heart sound by the auscultation technique just adds on to the analysis of our results by stating the mechanical conditions of the Heart [9, 10].

2 Methodology

A detailed study was carried out on the need for a low-cost device which would detect subdued abnormalities related to cardiovascular disorders. The device when coupled with any low-cost user-friendly mobile phone would analyze and record the data in real time and alert the patient when necessary.

Design of the device was done in two phases. The hardware implementation (Phase I) and software implementation (Phase II). Upon successful development of Phase I and testing of the device, the data was captured. This data was fed to the software module in Phase II for further processing. Hardware implementation (Phase I) comprised device development and testing. In this phase, after design and development of the Hardware component, we used our indigenous device to acquire the ECG signals from consenting volunteers in different age groups. This would at the same time test the hardware design and enable to collect data in real time (Fig. 2).

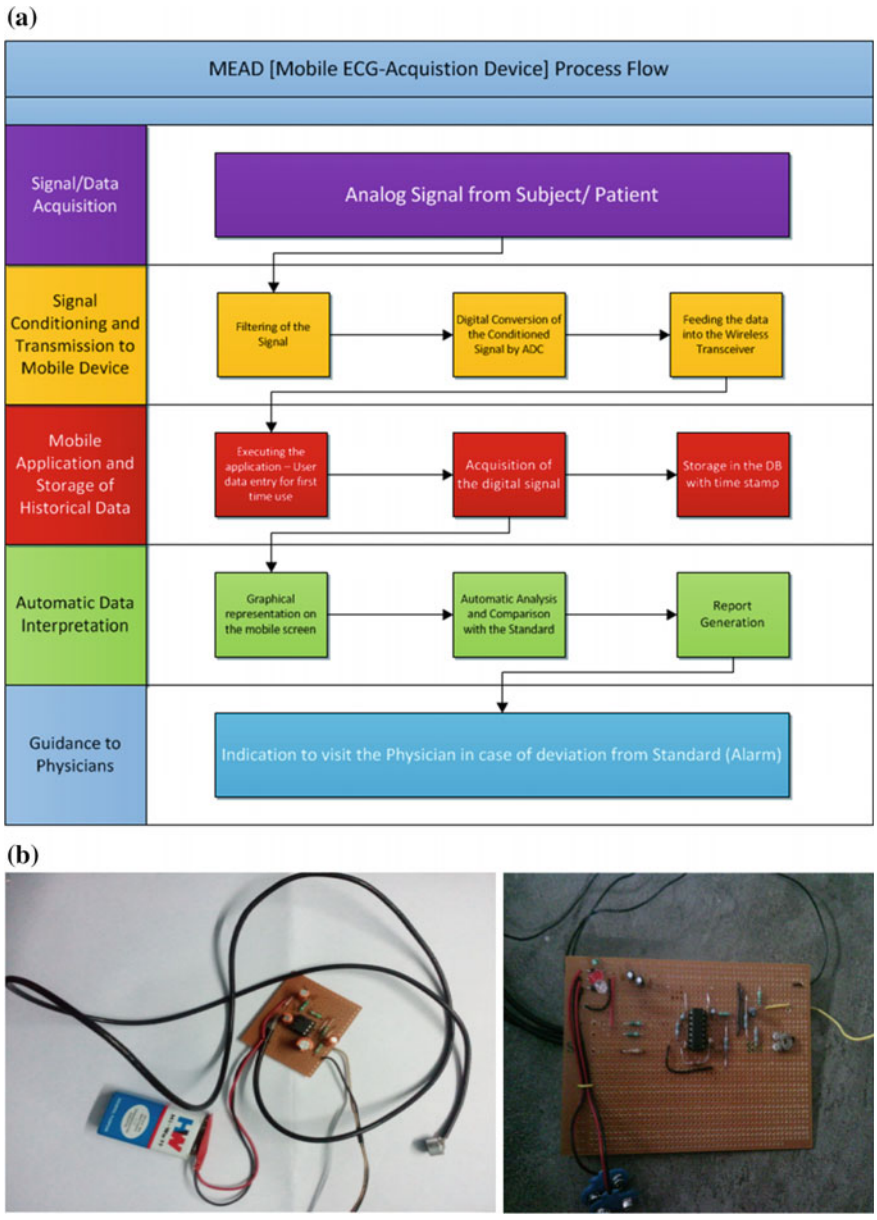


Fig. 2 **a** Design of CHMS. **b** Indigenous low-cost CHMS circuit board. **c** Testing carried out on consenting adult volunteers

Development of the CHMS was done using indigenously available components in our laboratory. Once clear and amplified signals were achieved the need was to put the device to test. For this purpose, all our volunteers chosen were consenting

Fig. 2 (continued)

adults within the age-group of 18–75 years of age (both male and female). All volunteers were in cautious state of mind and performed normal activities like walking, standing or mild jogging. Our objective was to collect enough data to perform statistical analysis to determine the accuracy and significance of the data acquired by our CHMS.

3 Results

The graphs obtained from the custom CHMS for each of the cardiac signals are shown. The picture clearly depicts the changes in the waveforms for different physiological conditions.

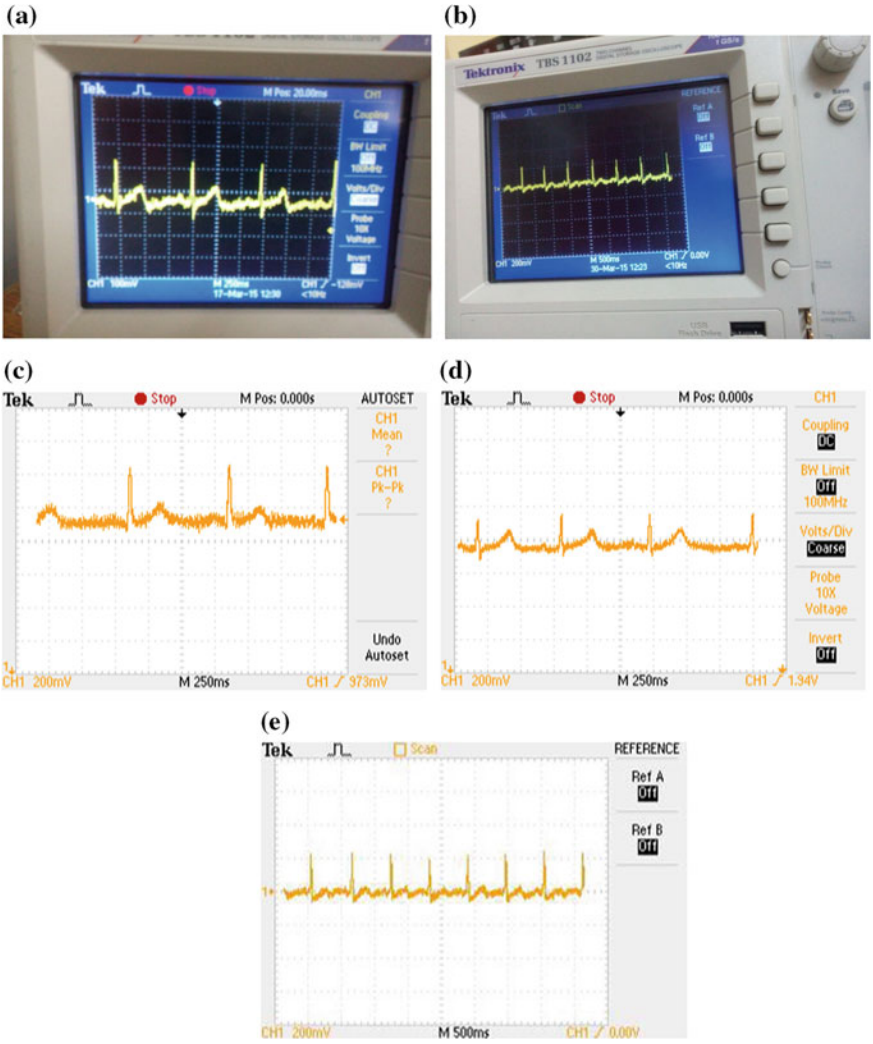


Fig. 3 **a** ECG signal at rest. **b** ECG signal after jogging. **c** Abnormalities in ECG in different subjects. **d** PPG signal recorded by CHMS. **e** ECG signal recorded by CHMS

3.1 ECG Signal Recorded Using CHMS

The following figures depict ECG Signals showing the different types of abnormalities like arrhythmia, elevated T-wave, diminished P-wave(atrial fibrillation) in different subjects.

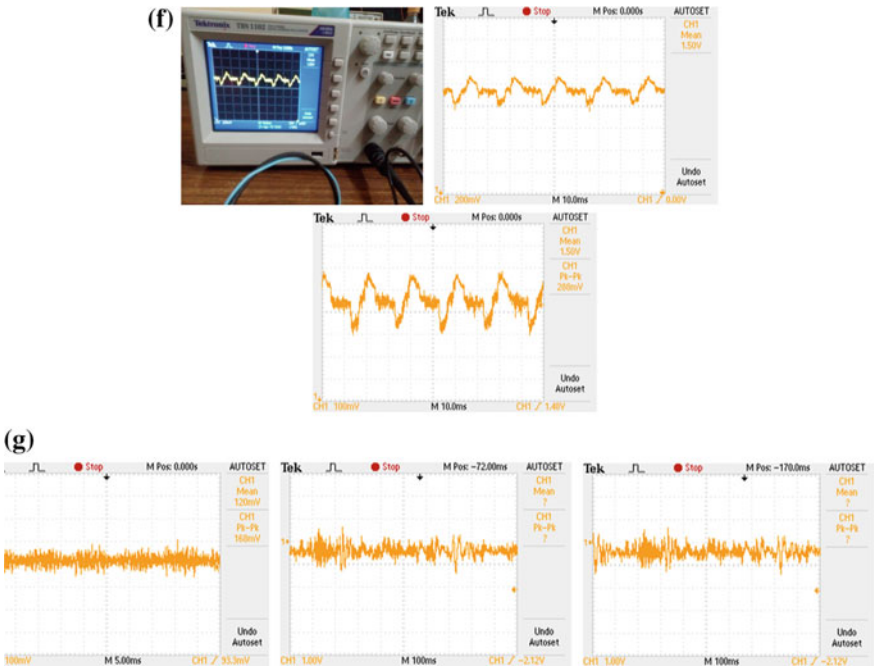


Fig. 3 (continued)

3.2 Photoplethysmography Signal Obtained Using CHMS

See Fig. 3.

3.3 Heart Sound Signal Recorded CHMS

From the obtained graph, it is clear that the original signals gets suppressed due to the power line and other interferences from mobile phones, laptops, watches, etc. which makes it difficult to count the heart beat from the heart sound signal.

4 Data Analysis and Interpretation

The following table shows the different parameters that are retrieved from an ECG Signal of different Subjects who were all adult consenting volunteers under various conditions, as recorded from the custom-made indigenous device (Fig. 4 and Table 1).

Simple linear regression was conducted with *R-R interval* as the independent variable and *heartbeat* as the dependent variable. The objective of doing this

analysis was to test whether *R-R interval* has a significant effect on *heartbeat*. There were 21 complete readings. Results show that there is a significant effect of *R-R interval* on *heartbeat* ($\beta_{unstandardized} = -96.5$, $p < 0.05$). One unit increase in R-R interval causes a decrease in heartbeat by 96.5 units.

Coefficients[a]						
Model		Unstandardized coefficients		Standardized coefficients	t	Sig. (p-value)
		B	Std. error	Beta		
1	(Constant)	152.355	1.823		83.596	0.000
	R-R Interval	−96.510	2.199	−0.995	−43.894	0.000

[a]Dependent Variable: Heartbeat

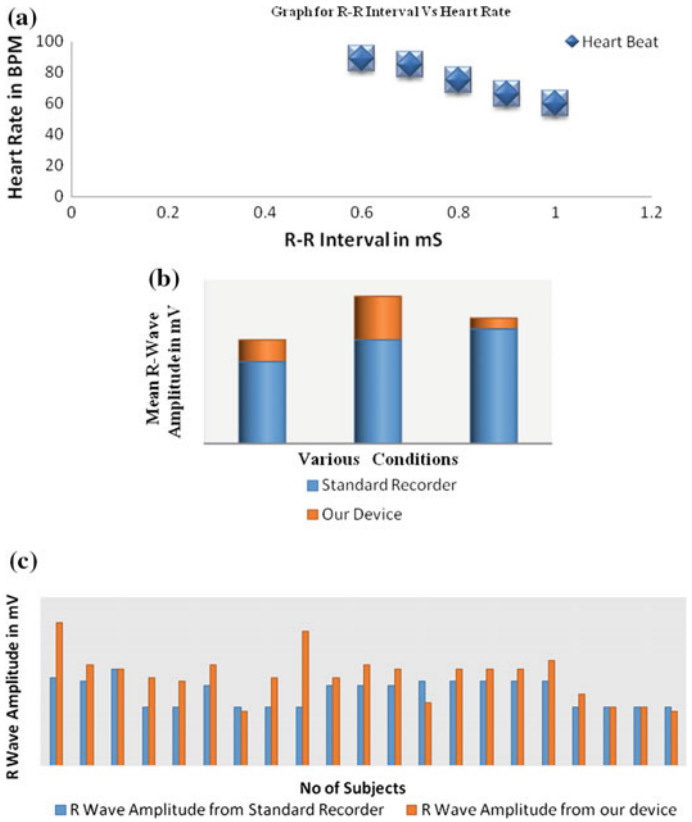
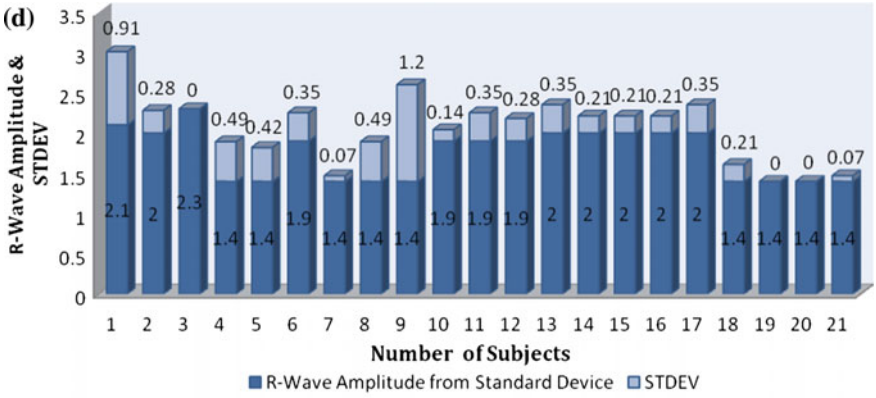
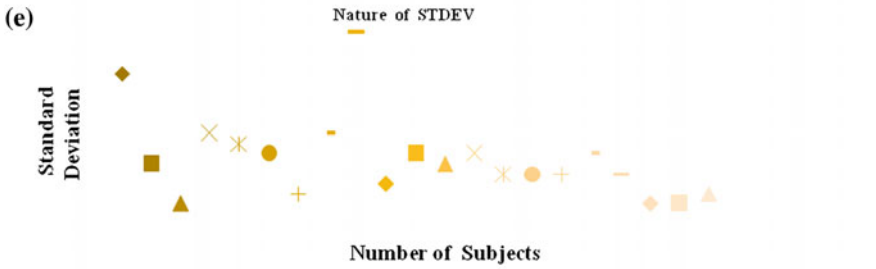


Fig. 4 **a** Plotting of R-R interval and corresponding heart rate. **b** Plotting of mean R-wave amplitude during different conditions recorded by the standard device and the custom-made device. **c** Plotting of R-wave amplitude recorded by the standard and custom-made device. **d** Plotting of standard deviation of R-wave amplitude. **e** Plotting of the nature of the standard deviation. **f** Plotting of comparison of pulse rate





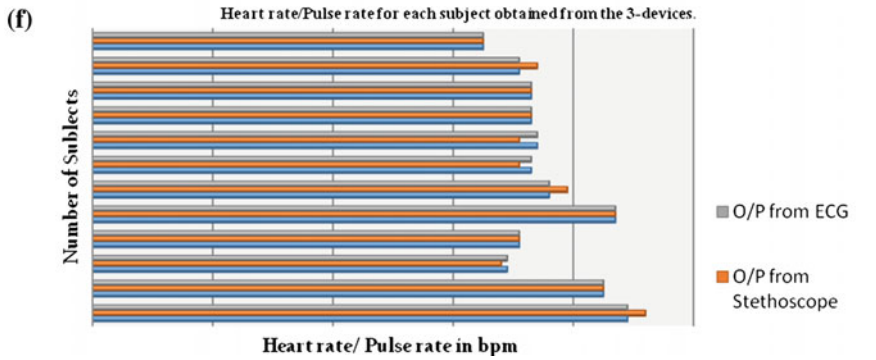


Fig. 4 (continued)

Table 1 Parameters retrieved from ECG signals recorded by our CHMS

Sl. no.	Subjects[a]	Age (Yrs)	Gender	ECG signal parameters			
				Amplitude (mV)	Frequency (Hz)	R-R interval (ms)	Heart rate (beats/min)
1	Subject 1	19	M	3.4	106.45	0.67	89
2	Subject 2	20	M	2.4	90.8	0.70	85
3	Subject 3	20	M	2.3	49.38	0.86	69
4	Subject 4	21	F	2.1	52.56	0.86	69
5	Subject 5	21	F	2.0	55.68	0.84	71
6	Subject 6	21	M	2.4	98.07	0.68	87
7	Subject 7	22	F	1.3	55.64	0.86	69
8	Subject 8	24	F	2.1	67.89	0.89	67
9	Subject 9	24	F	3.2	44.57	0.92	65
10	Subject 10	24	F	2.1	152.8	0.82	73
11	Subject 11	25	F	2.4	56.45	0.86	69
12	Subject 12	25	F	2.3	51.8	0.78	76
13	Subject 13	29	M	1.5	89.8	0.82	73
14	Subject 14	36	M	2.3	42.08	0.82	73
15	Subject 15	38	M	2.3	46.9	0.86	69
16	Subject 16	42	F	2.3	78.6	0.81	74
17	Subject 17	53	F	2.5	59.45	0.82	73
18	Subject 18	55	M	1.7	62.3	0.86	69
19	Subject 19	62	M	1.4	66.45	0.86	69
20	Subject 20	69	M	1.4	60.8	0.84	71
21	Subject 21	74	M	1.3	39.38	0.92	65

The amplitude of the "R" wave of the ECG signal of different SUBJECTS under different conditions that have been recorded by the standard ECG recorder available in the market and the custom-made device are given below in the following table. Thus the standard deviation for the amplitude of R-wave under various conditions is calculated and plotted accordingly (Table 2).

One-way ANOVA was conducted with *type of device* as the as the independent variable and *R-wave amplitude* as the dependent variable. The objective of doing this analysis was to test whether *type of device* has a significant effect on *R-wave Amplitude*. The sample size was 42. Results show that there is a significant effect of *type of device* on *R-wave Amplitude* ($\text{Mean}_{\text{our device}} = 2.12$, $\text{Standard Deviation}_{\text{our device}} = 0.56$ versus $\text{Mean}_{\text{standard device}} = 1.74$, $\text{Standard Deviation}_{\text{standard device}} = 0.31$, $F(1, 40) = 7.53$, $p < 0.05$) (Tables 3 and 4).

Table 2 Calculation of standard deviation for the R-wave amplitude

Sl. no.	Subject[a]	Condition	Amplitude of R-wave from standard recorder (mV)	Amplitude of R-wave from custom-made device (mV)	Standard deviation (SD)
1	Subject 1	Running	2.1	3.4	0.91
2	Subject 2	Jogging	2.0	2.4	0.28
3	Subject 3	Running	2.3	2.3	0
4	Subject 4	Normal	1.4	2.1	0.49
5	Subject 5	Normal	1.4	2.0	0.42
6	Subject 6	Staircase	1.9	2.4	0.35
7	Subject 7	Normal	1.4	1.3	0.07
8	Subject 8	Normal	1.4	2.1	0.49
9	Subject 9	Normal	1.4	3.2	1.2
10	Subject 10	Normal	1.9	2.1	0.14
11	Subject 11	Staircases	1.9	2.4	0.35
12	Subject 12	Staircases	1.9	2.3	0.28
13	Subject 13	Normal	2.0	1.5	0.35
14	Subject 14	Brisk walking	2.0	2.3	0.21
15	Subject 15	Brisk walking	2.0	2.3	0.21
16	Subject 16	Staircases	2.0	2.3	0.21
17	Subject 17	Staircases	2.0	2.5	0.35
18	Subject 18	Normal	1.4	1.7	0.21
19	Subject 19	Normal	1.4	1.4	0
20	Subject 20	Normal	1.4	1.4	0
21	Subject 21	Normal	1.4	1.3	0.07

[a]All Subjects were adult consenting volunteers

Table 3 Descriptive Statistics

Dependent variable: R-wave amplitude			
Type of device	Mean	Std. deviation	N
Standard recorder	1.7429	0.31555	21
CHMS	2.1286	0.56138	21
Total	1.9357	0.49031	42

The pulse rate or the heartrate (in bpm, as both are same) of different subjects is simultaneously recorded by the designed circuit for photoplethysmography, calculating the R-R interval from ECG and manually by the stethoscope (Table 5).

Table 4 Tests of between-subjects effects

Dependent variable: R-wave amplitude							
Source	Type III sum of squares	df	Mean square	F	Sig. (p-value)	Noncent. parameter	Observed power[b]
Corrected model	1.562[a]	1	1.562	7.534	0.009	7.534	764
Intercept	157.374	1	157.374	758.949	0.000	758.949	1.000
Type of device	1.562	1	1.562	7.534	0.009	7.534	.764
Error	8.294	40	0.207	–	–	–	–
Total	167.230	42	–	–	–	–	–
Corrected total	9.856	41	–	–	–	–	–

Table 5 Comparison of pulse rate obtained from different recordings

Sl. no.	Subject	Pulse rate from custom-made device	Heart rate from the stethoscope	Heart rate calculated from R-R interval
1	Subject 1	89	92	89
2	Subject 2	85	85	85
3	Subject 3	69	68	69
4	Subject 4	71	71	71
5	Subject 5	87	87	87
6	Subject 6	76	79	76
7	Subject 7	73	71	73
8	Subject 8	74	71	74
9	Subject 9	73	73	73
10	Subject 10	73	73	73
11	Subject 11	71	74	71
12	Subject 12	65	65	65

5 Conclusion

Comparative study and analysis of results obtained by our CHMS showed that the data obtained from different parameters are significantly close to the data obtained from expensive instruments available in the market. Diagnostic significance of the data as interpreted by a physician provides a similar conclusion as obtained from standard device available in the market which are often expensive and difficult to afford by people of low-income groups dwelling in rural areas. Conduction of ANOVAs on R-wave amplitude recorded by our CHMS is statistically significant and can be used as a tool for abnormalities associated with R-wave disorder. Ventricular disorders are commonly associated with changes in R-R interval and R-wave amplitude hence can be easily diagnosed by our CHMS. Considering the low-cost and indigenous components used which makes troubleshooting lot easier our CHMS can been achieved. Our future work is intended to get a clearer heart

sound signal in real time. We will be focused in interfacing our device with the common handheld phones so that the data can be stored using the phone memory for easy retrieval and reference as and when required. The digitized data acquired by the CHMS will be transferred to a smart handheld mobile device. The excess power will be utilized for further processing which includes visualizing the data in a readable format.

Acknowledgements This work has been partially funded by UGC Minor Govt. of India.

References

1. American Heart Association: Heart disease and stroke statistics-2008 Update. AHA, Dallas, Texas (2008)
2. Prevention of Cardiovascular Disease Guidelines for Assessment and Management of Cardiovascular Risk. World Health Organization (2007)
3. Dutta, M., Dutta, S., Sikdar, S., Dutta, D., Sharma, G., Sharma, A.: Reliable, real-time, low cost cardiac health monitoring system for affordable patient care. In: Lecture Notes in Bioengineering: Advancements of Medical Electronics. Springer, India (2015). ISBN: 978-81-322-2255-2
4. Goldberger, A.L., Goldberger, E.: Clinical Electrocardiography: A Simplified Approach, 7th edn. Elsevier/Mosby Inc, St Louis (2006)
5. Fox, et al.: Resting heart rate in cardiovascular disease. J. Am. Coll. Cardiol. **50**, 823–830 (2007)
6. Amin, D.S.M., Fethi, B.R.: Features for heartbeat sound signal normal and pathological. Recent Pat. Comput. Sci. **1**, 1–8 (2008)
7. He, B., Cohen, R.J.: Body surface Laplacian ECG mapping. IEEE Trans. Biomed. Eng. **39** (11), 1179–1191 (1992)
8. Oresko, J., Cheng, J., Jin, Z., Duschl, H., Cheng, A.: Detecting cardiovascular diseases via real-time electrocardiogram processing on a smart phone. In: Proceedings of the Workshop on Biomedicine in Computing: Systems, Architectures, and Circuits (BiC), pp. 13–16, June 2009, Austin, TX
9. Bhatikar, S.R., DeGroff, C., Mahajan, R.L.: A classifier based on the artificial neural network approach for cardiologic auscultation in pediatrics. Artif. Intell. Med. **33**(3), 251–260 (2005)
10. Zhang, G.Q., Zhang, W.: Ageing Res. Rev. **8**(1), 52–60 (2009)

Part X
Outcome Based Education

Some Limitations of Outcome-Based Education

Avik Sanyal and Rajashree Gupta

Abstract Outcome-Based Education or OBE as in short we know, means organizing for results: based on what we do instructionally on the outcomes we want to achieve. Outcome-based practitioners start by determining the knowledge, competencies, and qualities they want students to be able to demonstrate when they finish school and face the challenges and opportunities of the adult world. OBE, therefore, is not a "program" but a way of designing, delivering, and documenting instruction in terms of its intended goals and outcomes. During the 1980s Americans concluded that their schools were in serious trouble and that many children were not learning. In fact, children learn all the time; they are wonderful learners. They learn how to talk and walk, ride bicycles, and handle their parents and teachers. Schools take on the responsibility for planning the learning of children through the curriculum. As numerous educational reports and studies have recently shown, schools do not do this very successfully. One of the panaceas offered as a solution to this problem is an optimistic philosophy rooted in a "success for all" ideology of educational restructuring, a philosophy that may be broadly called "Outcome-Based Education" (OBE). The following research paper contains the limitations of Outcome-Based Education.

1 Introduction

"Education is a process, not a product" wrote Jerome Bruner 30 years ago. That statement is as true now as it was then. Education is a social-reflexive process that must be negotiated in classrooms on a daily basis.

A. Sanyal (✉)
Centre for Management Studies, JIS College of Engineering, Kalyani, Nadia 741235, India
e-mail: aviksanyal24@gmail.com

R. Gupta
Shantipur B.Ed College, Shantipur, Nadia 741404, India
e-mail: rajashree.munmun@gmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_57

Can a rational alternative to OBE be developed? Criticizing this model without suggesting a suitable alternative vision for curriculum and instruction is inappropriate. My purpose is not only to critique with a view to inviting experimental testing that would lead to refutation or refinement, but to posit an alternative "procedural-inquiry" model of education. Thus, I am critically analyzing the Outcome-based education.

I contend that curriculum and instruction can be effectively organized by a logic other than the "technical-rational planning" of outcomes. Clearly, OBE is a "means-ends" model of curriculum planning.

OBE means organizing for results: basing what we do instructionally on the outcomes we want to achieve… Outcome-based experts start by shaping the knowledge, skills, competencies, and qualities they want their students to demonstrate when they finish their individual school and be able to face the world.

Another supporter of OBE has argued that in a learning-centered classroom the focus is on the outcomes, not the methods, materials, or, indeed, evens the learning experiences. Outcome-based education can be described as a "training-instructional" model that views schools as vocational skills dissemination centers and not educational environments.

One approach to the problem presented by OBE is to suggest the use of different planning models, depending on whether those areas are based on knowledge, training, instruction, skills, and so on. For example, OBE may be highly suitable for teaching technical writing skills; but the teaching of art or English literature does not take the form of a step-by-step progression toward outcomes. If our aim is to get pupils to use knowledge creatively, then it is nonsensical to try to define specific behaviors that will result from education. The crucial thing about knowledge is that we *think* with it. How can I, as a teacher of English literature, define what a student will have as an outcome from reading *Paradise Lost*?

Claims that OBE can be applied to the entire curriculum presuppose that objectives are appropriate for all subjects, at all levels of education. Teachers and theorists in the arts and humanities in particular have countered that in these fields the concern is not for students to reach goals or exit outcomes once and for all, but rather for them to develop standards of judgment, criticism, and taste.

Spady and Marshall in their book "Beyond Traditional Outcome Based Education", couch their arguments for OBE in a "success for all" vision. They contend that:

- All students can learn and succeed.
- Success in school breeds further success.
- Schools control the conditions of success.

Well, this may be true for some pupils and some subjects. But perhaps this theory needs refinement-or even rejection. The idea of significant, observable educational outcomes and the notion of curriculum as preparation for adult life is not new; the "objectives model" formulated by educators such as Franklin Bobbitt and Ralph Tyler in their book "The Curriculum" (Boston: Houghton Mifflin, 1918); and Basic Principles of Curriculum and Instruction (Chicago: The University of Chicago Press, 1949) exalted the instructional objective.

1.1 Objections to Outcome-Based Education

The most fundamental criticism against OBE is that it minimizes education, teaching, and learning to forms of human engineering and quasi-scientific planning procedures-procedures that view education as an instrumental means to specified ends. This model, which many educators find unacceptable, amounts to molding students through behavior modification. It resembles the activity analyses of human behavior discerned by Bobbitt in his book "The Curriculum", 75 years ago.

The educated mind will always achieve unique and novel interpretations because knowledge is a tool to think with. To cite the significant outcomes in advance of teaching and learning is absurd.

A second objection to OBE relates that knowledge and set of courses content can be sequentially broken down into "small-outcomes" that in due course lead to more significant "exit outcomes." This view of knowledge disregards the theory of knowledge. Translation of the deep structure of knowledge into simple outcomes is, at minimum, a gross distortion. Knowledge and understanding can never be reduced to behaviors, lists of skills, and observable performances. Knowledge is an open-ended inquiry, not some product, or outcome to ultimately reach. To adopt the OBE stance is to trivialize knowledge, to reduce it to objective facts.

A related problem is the implication that there are systematic hierarchies of objectives, beginning with lessons and continuing through units, courses, programs, and, indeed, the entire educational career of the student. This linear, step-by-step view is a little too tidy for most learning that goes on in schools. We do not acquire factual knowledge that automatically leads to understanding. In my view, knowledge and accepting and affect are confluent phenomena that go hand-in-hand in some cases. True learning, like excellent teaching, is more arguably developmental and not linear.

Thirdly, while OBE may improve the structure of lessons and units within courses, it does not necessarily improve the quality of curriculum. We have no empirical evidence that OBE functions more effectively than a "process" model. No evidence suggests that this is how pupils best learn or understand their educational experiences. If 90% of our students attain honors grades in trivial pursuits and experiences, curriculum and learning are not advanced. Moreover, teachers, policy makers, and others with an educational stake do not have a tradition of teaching or planning in this outcomes-directed way.

A fourth objection relates to the inclination for outcomes to be expressed as simple "recall" or "learning" objectives, often because these are easiest to assess. This tendency helps explain the enormous drive toward centralized state assessment procedures. In fact, OBE is characterized by state prescriptions of student outcomes measurable by external testing. This characteristic diminishes the professionalism of teachers by reducing their involvement in research and assessment activities.

One could argue that it is arrogant to suggest that outcomes, as expressed in paper and pencil assessments or examinations, define knowledge of a field, or of a student. More often assessments or examinations determine what the student has

not learned. Knowledge has more in common with speculation than with mastery. In addition, many significant outcomes may express themselves only over the long run or in the face of particular contingencies. For example, I worked on constructing a cultural studies program in Northern Ireland that aimed at encouraging "mutual understanding and tolerance in intergroup relations." Such a significant and important goal cannot be realized (or assessed) in a unit or term, but may take many years to manifest itself.

Fifth, a linear OBE model is non-reflexive-it does not examine itself. It specifies the limits of the field of study and suggests that there cannot be more education than is encompassed by the stated outcomes. In this sense it is really a ″minimalist″ instruction model, and furthermore, it implies a poverty-stricken model of student–teacher interaction. To see education as an instrumental means to predetermined outcomes places the teacher uncomfortably close to playing the role of indoctrinator. Indeed, one could argue that it is undemocratic to dictate the ends of learning in advance of teaching and learning. Education requires inquiry, and we cannot predict the ends of pure inquiry. "How can you put on the blackboard the mysterious internal goal of each creative person?''

This is not to say that we should not have an aim, or outcome. To teach without a goal would be irrational. But OBE carries much of the baggage of behavioristic psychology applied to schooling and curriculum. Its basic argument suggests that education should be about planning behavioral changes in students' performances. It further brazenly argues that all of the significant outcomes can be specified in advance. Moreover, it implies that all pupils should demonstrate similar outcomes and behaviors. I must ask, is education about some standard packaging of outcomes as products, or is it more akin to a reflective social process?

2 An Alternative: The Procedural-Inquiry Model

As indicated above, the outcome-based approach may be satisfactory for areas of training, demonstration, and low-level skills such as those required in vocational courses; but it clearly breaks down in the arts and humanities, where there is a trenchant concern for using knowledge to produce meaning.

In certain areas of the curriculum—those focusing upon knowledge and understanding—an alternative to OBE, which I shall call a "procedural-inquiry model," is more appropriate. The great advantage of the procedural model is that it rests ultimately on the strength of the teacher. Table 1 lists defining characteristics of the OBE and procedural-inquiry models.

The procedural-inquiry model starts not with the specification of ends or outcomes, but with the principles of procedure for doing inquiry in a particular field or form of knowledge. It does not presuppose some lockstep, linear progression through a continuum of goals from the level of lesson on through unit, course, program, and finally, national agenda. It is about teaching through inquiry, and

Table 1 Characteristics of outcome-based and procedural-inquiry models of education

Outcome-based model	Procedural-inquiry model
Exit outcomes; unit and lesson objectives	Broad curriculum aim stated in terms of pupil understanding and knowledge
Standards of student performance embodying the goals and objectives	Consent selection based up on selection of materials, methods and concepts representing criteria and procedures for doing the subject
Content selection based upon goals and objectives	Unit method not necessary
Materials and units sequenced in logical fashion	Teaching viewed as reflective social practice
Curriculum divided into micro-units	Inquiry-discovery orientation as instruction-training
Instruction directed as specified goals	Emphasis on divergence, initiative thinking, and quality of experience
Convergence emphasized	Many possible outcomes
Time adjustment for mastery depending on student aptitude	Complex evaluation with teacher as judge-researcher
Feedback and correctives given to sonmasters	Qualitative and quantitative data collected
Assessment and evaluation through objective tests	Assessment judgemental and interpretive
Assessment as measurement according to mastery standard	

Note In OBE, teaching and content are selected to attain pre-specified objectives. See James F. Block, H.E. Efthim, and R.B. Burns, *Building Effective Mastery Learning Schools* (New York: Longman, 1989). In the procedural-inquiry model curriculum knowledge is selected because it personifies the key concepts and inquiry methods for proceeding in the field. These aspects are intrinsically important in their own right and not because the work leads to an extrinsic goal or objective. See Lawrence Stenhouse, *An Introduction to Curriculum Research and Development* (London: Heinemann Educational, 1975)

evaluating teaching and learning through teacher classroom research. As such, it is no doubt demanding, but the benefits in terms of teacher self-development are enormous.

The procedural-inquiry model has three parts: (1) a broad aim, (2) principles of procedure, and (3) criteria for assessing student work.

The broad aim of the procedural-inquiry model is to advance understanding of social situations and controversial issues and the human and moral values thrown up by these issues. The principles of procedure are as follows:

- Discussion is the best teaching strategy for enhancing understanding.
- The teacher remains "neutral" on moral value issues.
- The teacher adopts a facilitator role and "chairs" the discussions to ensure continuity, summary, and access to evidence.

The following criteria are used for assessing student work:

- To what extent can students use knowledge and concepts to explore issues? For example, in discussions, the students might be asked to use the concept of role to compare the situations of men and women. A subsequent examination might ask the students to define "role" and use the concept in discussing relations between the sexes.
- To what extent can students understand a wide range of views on an issue? For example, a discussion in which students consider many different points of view on marriage might lead to an examination in which students are asked to give two accounts on alternative forms of marriages.

In the arts and other disciplines we can construct curriculum, not from the outcomes, but from the "incomes"; the content can be selected, justified, and evaluated according to the built-in criteria of that particular form of knowledge. To be able to "do" mathematics, for example, one must have respect for four aspects of the epistemology of mathematics:

- A wide base of knowledge, characterized by a central logical structure
- A network of key concepts that give meaning to the discipline
- Recognized modes of inquiry that characterize how new knowledge is added to the field
- Tests of truth, or proof, that can be used to explore experience or test, empirically, expressions of knowledge and thereby disprove or prove knowledge of the field

It seems irrational to have specified objectives in areas of the curriculum that seek to elicit a creative or individual response from students-for example, music, art, poetry; how could we decide in advance what each student is to become?

However, curriculum consists of more than the disciplines of knowledge. There are fields of knowledge in the practical and theoretical realms, for example, geography and engineering. In a bold move, Stenhouse proved that the process model could be adapted to fit school subjects other than the formal disciplines of knowledge. In the Humanities Curriculum Project, an integrated high school curriculum, Stenhouse's team employed the process model, eschewing objectives, and cast the teacher in the role of a researcher committed to certain principles of procedure in teaching.

The model I pose as an alternative to OBE is a rational planning model for curriculum described as a procedural-inquiry model of education because its design does not have to be based upon the specification of outcome blueprints but on the employment of the "principles of procedure" in a particular field or form of knowledge. Within the curriculum, inquiry by students and teachers feeds knowledge about teaching and learning. Teachers are always testing their curriculum as a hypothesis. Does it work well, satisfactorily, or poorly? By working with principles of procedure teachers adopt a research, or inquiry, stance toward teaching, which asks for self-evaluation of professional development and judgment. Our curriculum

design stipulates a significant aim (for example, to advance understanding of key concepts in social studies). It then specifies the procedures to be followed by teachers and students in the educational activities; it attempts to use key concepts in questions and criteria by which to judge the work. Content may be selected to illustrate the best procedures, concepts, and criteria. In a fundamental sense, the teaching processes and principles become the "objectives" or outcomes. It is thus very much a pedagogical model to test teaching and learning. A style of pedagogy that adopts action research or action inquiry-studied enactment can provide a basis for the teacher to reconceptualize OBE into a theory of professional self-development, because the teacher's ideas are integral to professional and practical competence.

3 Teaching and Learning

Teaching and learning involve much more than a technical specification of content, and scope and sequence charts littered with "outcomes." We recognize that rational planning and the need for positive results are prized by Western cultures. The problem is that many curriculum designers can hardly envision life in a different way.

The first and last principle of teaching and learning should be the special caring relationship between tutor and student. The word *education* comes from the Latin *educare*, meaning *to rear*, just as a mother rears her children. Rearing in this sense connotes loving and caring. It is not some form of engineering one in particular directions. It is artistic and creative, because the student thinks and creates meaning. Learning is a form of "meaning making." *It is not the goal of teaching to produce results, but to create an experience in which the student can arrive at creative encounters, be drawn out, and make meaning*. Curriculum is not a thing to be "covered" by teachers; it is meant to create occasions in which learning takes place. I believe that teaching with effect requires an "I-Thou" relationship between the teacher and student. American educators have been denied what Europeans have called "pastoral care" in curriculum: systematic schemes whereby tutors take on a special role of caring for children and their learning. In fact, caring words such as *love*, *care*, and *caring* have been noticeably absent from the lexicon of public education in the country. Another feature of teaching and learning suggests that they cannot be directed at specific objective targets lest all freedom be lost. To learn one must be *free to learn*-not constrained by plans and objectives and various other obstacles. If the student is to have the freedom to learn, then teaching cannot be constantly directed toward various objectives. The technical OBE stance devalues professional judgment by not focusing upon the process of education but instead highlighting the ends or outcomes.

Other than this "procedural" principle, a second argument supports the notion of placing teachers in charge of assessment and evaluation: the teacher-as-inquirer is truly in charge of the classroom and is managing the learners and the learning.

Besides, classrooms are ideal laboratories for doing research that will enhance curriculum and educational theory. Hence, the international movement to involve teachers in research and learning through action.

4 Conclusions

OBE hints at "mastery," when in fact most teachers are learners along with their students. Teachers should not set themselves up as experts but as models of inquiry. Much of the argumentation for OBE has not come from teachers or curriculum designers but from those within the field of assessment and testing, and its lore has been widely adopted by central offices and educational policy makers. Outcome-based education is also objectionable from an ethical-deliberative point of view. It begins with outcomes and results and then goes to extremes to plan and deliver instruction that will mold and change students to become what we want them to be. It is a form of human engineering, not a process of education. There is no question that curriculum must have a goal. If a teacher works toward a long-term aim, such as to promote tolerance among students, and then specifies a set of procedures that work toward achieving that goal, states what content is to be used, elucidates the sort of classroom experiences necessary, and builds in evaluation of this experimental process, then she will of course be acting rationally and planning rationally. This process casts the teacher in the role of a researcher by examining practice as problematic and curriculum proposals as hypotheses. The procedural-inquiry model involves teachers as participants in curriculum planning, implementation, and research, taking a close look at students' and teachers' work as the basis for school and curriculum improvement.

A final cautionary sounding bell must be rung: OBE suits the technical rationality currently prevailing in the United States and other Western nations whose policies emphasize high-tech culture and the preparation of students to compete in the workplace for global economic warfare. This skills-oriented model views schools as vocational centers producing workers and rests upon the argument that skill requirements on the job change faster than do curriculum and organizational changes in schools.

This skills model is not an artistic or creative response to the culture of schooling. It is an industrial model that views students as raw materials. As such, it explains the following statement issued by the U.S. Department of Defense Dependents schools: Based upon an assessment of the future we believe our students will face challenges and opportunities in a world characterized by worldwide economic competition and interdependence which creates ever-increasing requirements for job-related performance… A similar statement emerged from the Aurora Schools in 1991.

We will know we are accomplishing our mission when all our students are collaborative workers… and quality producers. We must not allow the education of the youths to continue to be dominated by the sort of industrial educational,

technological planning first presented by Bobbitt and later promoted by Tyler, Bloom, and a cavalcade of their students of mastery and OBE learning. Their values are clear: education as economic preparation and utility as the sole concern. No, we must be also concerned with valuing things for their own sake, not because they lead to some outcome.

Activity-Based Learning (ABL) for Engaging Engineering Students

Aniruddha Biswas, Sumit Das and Suchandan Ganguly

Abstract Activity-based learning is becoming an urge for modern day's learner centric, outcome-based education system. These techniques have been adopted within the Faculties of Engineering degree. This paper defines the adoption and application of these techniques. The paper reports on the design, development, and implementation of subject and discusses how it uses activity-based learning to ensure that students become more aware of design and team processes.

Keywords ABL—Activity-Based Learning · OBE—Outcome-Based Education · PLO—Program Learning Outcome · PEO—Program Educational Objective · CLO—Course Learning Outcome

1 Introduction

In the fast evolving twenty-first Century, Technology changes so frequently, thus education system is becoming lame to produce good quality Engineers. This phenomenon is very common in almost all the Engineering Colleges in India. To get rid of this unwilling situation most of the leading institutions have started many good practices and taken corrective measures. But it becomes a mandate to put all national level colleges under one umbrella. To make it happen, Govt. of India along with various Accreditation Committee has decided to standardize the Engineering Education System. Standardization is a huge responsibility in such a big country, where thousands of Engineering Colleges are there. And all these colleges are from

A. Biswas (✉) · S. Das · S. Ganguly
Information Technology, JIS College of Engineering,
Kalyani, Nadia 741235, West Bengal, India
e-mail: biswas.aniruddha@gmail.com

S. Das
e-mail: sumit.it81@gmail.com

S. Ganguly
e-mail: ganguly.suchandan@rediffmail.com

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_58

various states whose culture, weather; political, and socioeconomic structure is diversified. To create unity in diversity many Accreditation Committee like ABET, NBA, NAAC has introduced Outcome-Based Education System by following Washington Accord, a Global Standard. ABET and NBA has defined 11–12 Graduate Attributes for any Engineering Graduates. In OBE these attributes are kept in mind while defining the Program Learning Outcomes. In OBE system, Program Outcome is given highest priority. To achieve the outcome at the end of the program we need to change the Teacher-Centric traditional educational system to Student or Learner-Centric System. To produce a learner-centric environment, ABL is one of the best ways to follow.

2 Traditional Teaching Versus ABL

Traditional Teaching means Teacher-centric approach. In this environment only teacher plays an active role where as learners are passive mostly. But this technique is best suited for subject like "History" where knowledge transferring and remembering is the ultimate goal. Traditional Teaching mostly deals with unidirectional transferring of knowledge. Learners are mostly inactive, sometimes busy in taking notes but actual knowledge transfer does not occur [1]. Many learners are found to take wrong notes because of poor understanding of the lesson during class. So the effective learning does not take place which causes criticism of traditional teaching learning now-a-days. This criticism is made by OBE followers. But in Engineering it cannot help the students much to make them sufficiently efficient to fulfil the industry needs. Industry is demanding role ready graduates. But in many cases it is widely observed that even in campus recruitment industry is hiring very meager number of students who could actually fulfil their needs. The term called mass-recruitment is valid for very few Engineering Institutions. That is why a major change is required especially in Engineering Education. So that Graduates can fulfil the basic requirements of industry at least.

Here one quotable quote should be mentioned "Tell me and I forget, teach me and I may remember, involve me and I learn."—Benjamin Franklin.

Involvement of the students makes them learn quickly and easily. Activity-Based Learning adopted a range of pedagogical approaches in teaching. While the learner learns new things by engaging himself or doing hands-on experiments by his own, then learning climbs its highest peak.

The key concept of Activity-Based Learning comes from my classroom. Students in the same subject in two consecutive batches of first year are learning differently when I applied Teacher-Centric approach and Activity-Based Learning technique. If first year students are given ample chance to explore themselves then learning could be made easy, effective, and long-lasting.

Students frame questions, answer them, create new problem, solve them, think-pair-share, clears doubts, brainstorm in the class. Thus students are more involved than just listen; resulting higher learning curve (Fig. 1).

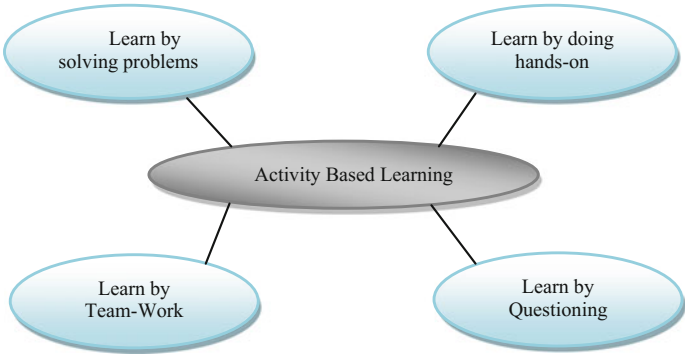


Fig. 1 ABL perspectives [2]

3 Paradigm Shift from Teacher-Centric to Student-Centric Approach

According to the traditional lecture-based format, where the students sit pseudo-actively and the professor delivers lecture, has become predominant in Engineering Institutions relatively recently. Earlier the passive students were provided with lecture notes that hardly help them to excel in that course. Now it is more important to engage them in some activity [3].

JISCE has supported a move toward more student-centric learning where professors are encouraged and facilitated in developing courses in such a way that the students could be engaged in deeper learning environment. The following principles are reflection of the above ideas in undergraduate education as described by Chickering and Gamson (1999, p. 76) [4]:

1. Student–Teacher Relationship
2. Student–Student Relationship
3. Active Participation
4. Feedback-based learning
5. Timeliness
6. Giving importance to diversified ways of learning.

The above principles are not that easy to be implemented as good practice in rural private engineering colleges as students are not so fluent in communication. To make it happen, we need to grow self-belief in students so that they can at least communicate with the faculty members as well as their classmates with full confidence. Thus students would become an active learner. Giving prompt feedback is must to the activities performed by the students. Positive feedback encourages students to achieve their goals. Thus students' participation and interaction is maximized.

4 Reason Behind This Shift

It has been observed that the effectiveness of Activity-Based Learning is much higher than that of traditional teacher-centric teaching methodology. In the first two years of undergraduate students, it has been seen that ABL is more fruitful approach than even Project-Based Learning. One or two decades ago it was observed that Engineering Students came from Govt. Schools where Traditional teaching was followed mostly. But recently it has been observed that Engineering students are coming from reputed private schools where activity-based learning are already practiced. So they obviously detest traditional teaching where teacher plays the key role and they become listener only. Now students are well prepared to do, act and play with activities. In this circumstance it becomes a mandate to shift the traditional teaching to activity-based learning in Engineering Education. Activity-based learning will help students to be proactive, focused, and creative.

Below are the key purposes of ABL:

- Give students the feeling of importance in the class room
- Give them small targets to achieve in class room
- Boost up their individual and team skills
- Increasing the thinking ability and creativity
- Teacher should play the role of facilitator to complete the given targets
- Building up communication skill among participants.

Listening to the teacher and replicating the same thing cannot be Engineering knowledge. Engineering knowledge refers to such kind of expertise that can help to solve varieties of similar kind of problem. And even can solve problems that may be inferred from that. Problem solving ability comes from doing activity. Activity must be started from the classroom itself. So to engage the students in doing activity and giving them tasks to perform will increase their critical thinking ability (Fig. 2).

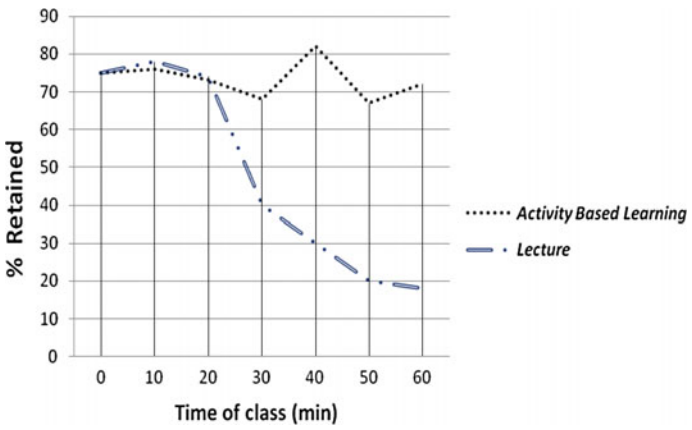


Fig. 2 Learning curve of lecture versus activity-based learning [2]

By this process they can analyze new things in a smart way and could lead to invention. Unless students start doing activity their knowledge cannot be tested and will be of no use. We all know that "Idle brain is Devil's Workshop." To get rid of that devil nature and mold it into a creative one, ABL can help engineering students to be empowered with ability to think, analyze, and create.

5 How ABL Helps Overall Growth of Students Learning Ability

In present technology driven society, information is available at the fingertips of the students. So only giving the lecture means repeating the same thing in same or little different way. That is why it is necessary to change the entire teaching learning process; so that the students' learning ability must increase. To achieve high learning ability, students need to be engaged in doing some activity. These activities will vary from student to student depending on his/her learning curve. But at every level teacher has to judge the student and give him/her a certain task in corresponding subject based on his/her competency. It will create a feeling to the students' mind that someone is there who is guiding to excel in his/her goals. Thus it will increase the student–teacher relationship. This can be best achieved if learners can be monitored closely by the teachers by engaging them with very small piece of activities depending on their capability. Thus ABL will help the students for better learning and also help to achieve Course Learning Outcomes (Fig. 3).

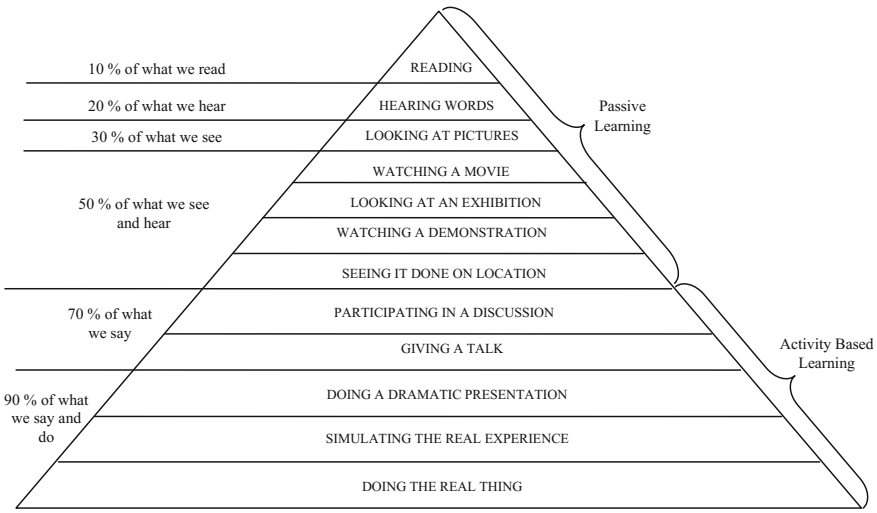


Fig. 3 After two weeks students tend to remember [2]

6 Contribution of ABL in Outcome-Based Education

Outcome-based education is a vast topic to discuss, in short it focuses on outcome. To achieve the outcomes objectives are defined. Objectives are theoretical whereas outcomes are practical. The ability to perform something by a graduate is known as outcome. To achieve the outcomes ABL can play the key role. Also helps to map the course learning outcomes and program outcomes in OBE system [3].

The problems we face in traditional teaching are that even if the syllabus is covered 100% but learning scale could be below 50% for more than 50% of students. So, this is really a harmful situation for coming future. Students will be graduated but without any expertise. This will be an unhealthy situation for the societal growth.

To recover from this unwilling condition corrective measures have been taken by many institutions. Among them OBE is leading the ladder. Each student should achieve Graduate Attributes set by various Accreditation Committee (e.g., NBA) if he/she is graduated from an institute which follows OBE. In OBE system graduates must possess 12 graduate attributes [5]. To achieve them many ways being followed, ABL is one the most effective one. How can we make ABL effective? This question can be answered as follows: First each and every course needs to be divided into modules. Proper lesson and session plan is to be made on each course. In each session what are the activities that can be assigned should be minutely specified at the highest possible depth. Thus different levels of students can be assigned with problems of different complexity levels (e.g., high, medium, low). Each students' progress can be measured by the teacher everyday at theoretical and lab classes.

Classification of ABL: list of strategies for active learning [6]:
Participation: Students are assigned with small tasks, so that they can be engaged in doing activities in class room.
Critical Thinking: Students task should be out of the box, not directly related to their base knowledge but can be answered after thinking.
Analyzing: Tasks given to the students must be subjective so that they can put their views in different ways after analysis.
Knowledge Sharing: Task given to the students must require varying knowledge, so that students are bound to communicate amongst themselves.
Team-work: After sharing of knowledge they can conclude and take decision from their team.
Communication: Team-work and knowledge sharing increase communication amongst team members. They also need to communicate with teacher to solve critical problems.
Problem Solving: Increase the ability to solve real life problems by analyzing and solving class-room problems based on course topics.
Quiz: It helps students to be more competitive in nature and answer quickly.
Debating: It increases the power of establishing the self argument in front of fellows.

7 Conclusion

A student can learn best only if he/she is involved as an active role player in the teaching learning process. Unless the habit of involvement is grown up in nature of the student, he/she cannot perform as a team lead in a group of people in some project. By adapting activity-based learning students can learn how to make mistakes and overcome from that. If student is never involved in solving the problem then he/she will never dare to do mistakes thus no new things will be invented if not tried at all. So thus we can conclude that ABL can be treated as the best approach to make the Engineering graduates capable of achieving all Program Learning Outcomes.

8 Future Scope

This paper could be used as a template for the design, development, and implementation of engineering subject and discusses how it uses activity-based learning.

Each subject could be planned at micro-level so that students can be engaged in teaching learning environment. In future, faculty members can work on different subjects in which they are expertized and make that ABL oriented. So that engineering students could be engaged in a learner-centric environment.

References

1. Maclellan, E., Soden, R.: The importance of epistemic cognition in student-centred learning. Instr. Sci. **32**(3), 253–268 (2004)
2. https://www.nescent.org/eog/documents/IntrotoActiveLearning.ppt
3. Fallon, E., Walsh, S., Prendergast, T.: An activity-based approach to the learning and teaching of research methods: measuring student engagement and learning. Irish J. Acad. Pract. **2**(1) (2013). Article 2. Available at: http://arrow.dit.ie/ijap/vol2/iss1/2
4. https://www.griffith.edu.au/__data/assets/pdf_file/0007/134278/EvaluatingSevenPrinciplesV2.pdf
5. http://files.eric.ed.gov/fulltext/ED380910.pdf
6. Gleason, B.L., Peeters, M.J., Resman-Targoff, B.H., Karr, S., McBane, S., Kelley, K., Denetclaw, T.H.: An active-learning strategies primer for achieving ability-based educational outcomes. Am. J. Pharm. Educ. **75**(9), 1–12 (2011)

Realization of Outcome-Based Education through Teaching Learning Methodology in Engineering Institution

Rajdeep Chowdhury, Soumyabrata Saha, Sudipta Sahana and Debashis Sanki

Abstract Although Outcome-Based Education has not been affluently realized and implemented in public education amid plentiful countries, it has been effectively espoused in the technical fields in India. Outcome-Based Education has become a substance of conversation in facade of public media, encompassing a tinge of squabble. Critics differentiate this modus operandi of edification in unenthusiastic terms. In the formulated paper, the authors have proposed an implementation modus operandi of Outcome-Based Education in the field of engineering and higher studies. The authors have embodied a specific correlation amid Institute vision, mission, Program Educational Objective, Program Outcome, and Course Objective and Course Outcome. Based on the Program Outcome, how curriculum and syllabus design and implementation would be adopted is represented in detail. Based on several assessments along with feedback mechanism, the entire procedure refinement would be accomplished that would lead to successful realization of Outcome-Based Education.

Keywords Outcome-Based Education · Program Educational Objective · Program Outcome · Course Objective · Course Outcome · Feedback

R. Chowdhury (✉)
Department of Computer Application, JIS College of Engineering,
Block–A, Phase–III, Kalyani, Nadia 741235, West Bengal, India
e-mail: dujon18@yahoo.co.in

S. Saha · D. Sanki
Department of Information Technology, JIS College of Engineering,
Block–A, Phase–III, Kalyani, Nadia 741235, West Bengal, India

S. Sahana
Department of Computer Science and Engineering, JIS College of Engineering,
Block–A, Phase–III, Kalyani, Nadia 741235, West Bengal, India

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9_59

1 Introduction

In this paper, realization of Outcome-Based Education through teaching learning methodology in engineering Institution has been substantiated with ample elucidation of the same. Academics have been strengthened over the years amid the unfathomable foundation of twin pillars, namely; Teaching and Learning. Teaching method comprises of the principles and methods employed for instruction to be implemented by facilitators to realize the desired learning by the students. The twin concept of Program Educational Objective and Program Outcome has been epitomized in its factual wisdom in the subsequent fragments and the manner in which they form an integral fraction of Outcome-Based Education embodies the crux of the paper.

Outcome-Based Education is a model of education that amend the traditional focus on existing teaching learning, in favor of making students exhibit what they know and are what they are able to do, irrespective of the requisite outcome. The emphasis in an OBE system is on measured outcomes rather than 'Inputs' and it reforms, emphasizing on setting apparent standard for visible and quantifiable outcome.

In traditional education approach, the system is essentially educator centric and outcome is not predefined. Traditional teaching is concerned with the teacher being the executive of the learning milieu. Authority and responsibility are seized by the teacher and they seize the role of the instructor (In form of lectures) and the decision maker (In respect to curriculum content and specific outcomes). They regard students having 'Knowledge holes' that need to be packed with information. In short, the traditional teacher observes themselves being the basis for learning to transpire. In Outcome-Based Education, initially the outcome is set and then to gratify the outcome, the objectives are set. Essentially traditional edification is top down approach whereas OBE is bottom up approach.

2 Related Work

There are visionary elements of Outcome-Based Education which some people find disconcerting. In [1] Spady proposed that the requisite for educational leaders to engage in empowerment thinking, visionary thinking and prospective concentrated thinking that gaze at the world, as it should be in the future. Spady criticizes 'Educentric Thinking' that focuses only on how the education system is at present. In an address to the Australian Primary Principals Association, Spady exemplified an advanced paradigm for educational reform [2] in terms like Essence of Model, Nature of Learning, Major Outcome Measure, Key Pedagogy, Temporal Structure, etc. In Outcome-Based Education, time permissible for realization of an outcome is assumed to vary amid students. Thus, the argument for achievement-based rather than time-based progression is ensured. Employing norm-referenced system of

grading and student comparative evaluation as prevailing form of assessment is discarded, posing an argument that the assurance of some students emerging from their schooling as failure, thus infringing Principle. Spady has been critical of employing National standardized testing which is visualized as closing off rather than expansion opportunity for students. Measuring student's performance against standards rather than merely against performance of other students is also debated in other educational circles. The implication of norm-referenced assessment for student's self-efficacy, goal attainment, and self-theories of intellect must be taken into account [3, 4]. The 'Forms of Knowledge' perspective is subject to imperative criticism such as it being narrowly focused on only intellectual objectives [5]. However, it is the prevailing approach employed in the specification of key learning areas, which endow with the basis of most curriculum frameworks in Australia and in other International countries.

Critics have found the language of Spady's vision for Outcome-Based Education problematic. Some outcomes are observed to lie outside the territory of the school, or to ensure privilege to some culture or perspective over others [6]. Others observe the outcome as indistinct and intricate to delineate in ways that make them useable by teachers [7, 8].

The Louden Report in WA reported that a substantial number of teachers expressed intricacy in translating expansive outcome statements in a pedagogically meaningful manner [9]. A novel academic Institution in the Gulf region have already tackled the above mentioned issue by adopting an academic framework that is based on Outcome-Based Education, while still employing the grade point average. The academic model is a hybrid approach that accommodates learning outcomes for measurement of the learning process and employs grades to accommodate the classic academic system. Anticipation ascertains the model would ensure that grade inflation is under control and students are attaining the learning outcome to become life-long learners [10].

3 Methodology: Process for Defining Vision and Mission

With respect to the Institutional mission and vision, Departmental mission and vision must be ensured. Nevertheless, after conduction of multiple meetings with different stake-holders, accumulating their valuable suggestions and feedbacks and finally on receipt of approval from Board of Studies (BOS), the student fraternity could cope up with prospective challenges in their profession and research work.

- **Program Educational Objective**

Defining Program Educational Objective would accentuate that they are broad statements that exemplifies the career and proficient accomplishments that the program is preparing graduates to realize. Program Educational Objectives play an imperative role as they embody the eventual mean to evaluate the eminence of a program (Fig. 1).

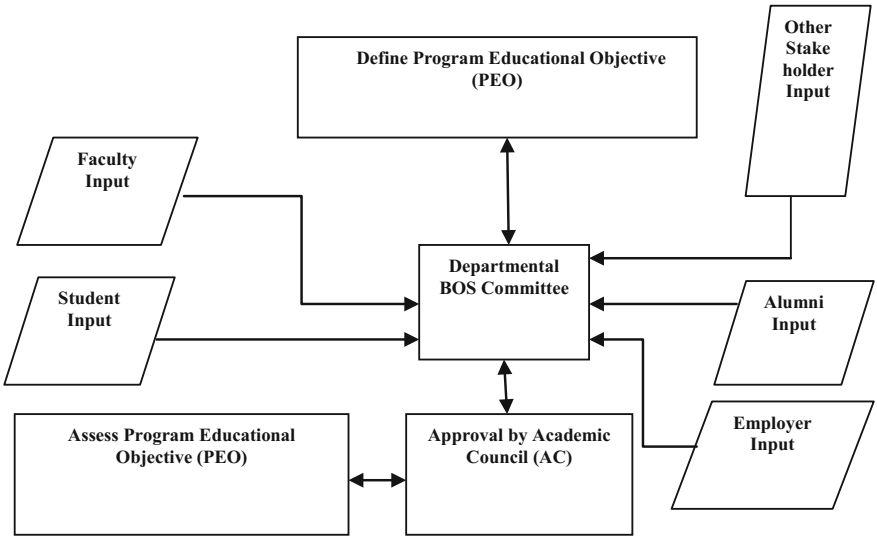


Fig. 1 Process employed for establishing the Program Educational Objective

- **Process employed for establishing the Program Educational Objectives**

Departmental mission is very much unswerving with the mission of the Institute. According to the Institutional mission, the Departmental mission is to be set. To attain the mission, PEO have been crafted to edify the student fraternity in such a way, that they could effortlessly cope up with any milieu. PEOs are the expansive statements of the objectives for which the program is run. In initial place, these objectives should aid in fulfilling the mission of the Department. Secondly, the students graduating from the Department are expected to lead an industrious and noteworthy life in the society, being useful in the progressive development, at all corners. Thus, it is essential that these objectives should be in consonance, to the extent possible, with the existing research scenario in the pertinent fields of engineering and industrial needs. The research scenario is best judged by the faculty members through research publications in an assortment of journals and conferences. The industry needs are gauged through the feedback mechanism, mostly verbal, received from the training and placement office, amassed during the company visit in campus recruitment drive. In addition, the industry scenario has a specific relationship with research and development; hence faculty members are somewhat proficient to assess the needs of industry. Moreover, the alumni who are situated in industry or Institutes of esteemed reputation are able to reflect whether the objectives are adequate.

• Responsibility of Administrative System in Ensuring the Achievement of PEOs

The three fold administrative system; Board of Governors—Academic Council—Departmental Committee, through effective planning and implementation, with self appraisal and feedback mechanism ensures the attainment of objective in a scalable/tangible manner. The Academic Council meeting at Institutional end and Departmental Committee in apiece Academic Department attempt to ensure unswerving Teaching Learning procedure.

In addition to this, Departmental representation in Institutional activities is ensured by the representation of faculty members from the Department in various committees. Departmental representation is visible in apiece statutory committee and cell of the Institution:

- Academic Council (AC)
- Board of Examination (BoE)
- Research and Development Committee (R&D)
- Monitoring and Evaluation Committee (MEC)
- Indent Scrutinizing Committee (ISC)
- Library Committee (LC)
- Sports Council (SC)
- Disciplinary Committee (DC)
- Different Committee Member under TEQIP
- IQAC and more

• Use of Diverse Tools and Processes in Assessment of Achievement of PEOs

Type of assessment tool	Assessment tool	Assessment criteria	Data collection frequency	Responsible entity
Direct	Assignment	Marks	Subject faculty	Subject faculty
	Class test	Marks	Twice a semester	Subject faculty
	Course performance	Number of students passed	Once apiece semester	Institute BoE
	Project	Marks in grade	Once apiece year	Institute BoE
Indirect	Placement record	Number of students placed	Once apiece year	Institute TPO
	Higher studies record	Number of students who opted for higher studies	Once apiece year	Department
	GATE score	Number of students with valid GATE score	Once apiece year	Department
	Alumni survey	Level of achievement	Once apiece year	Department

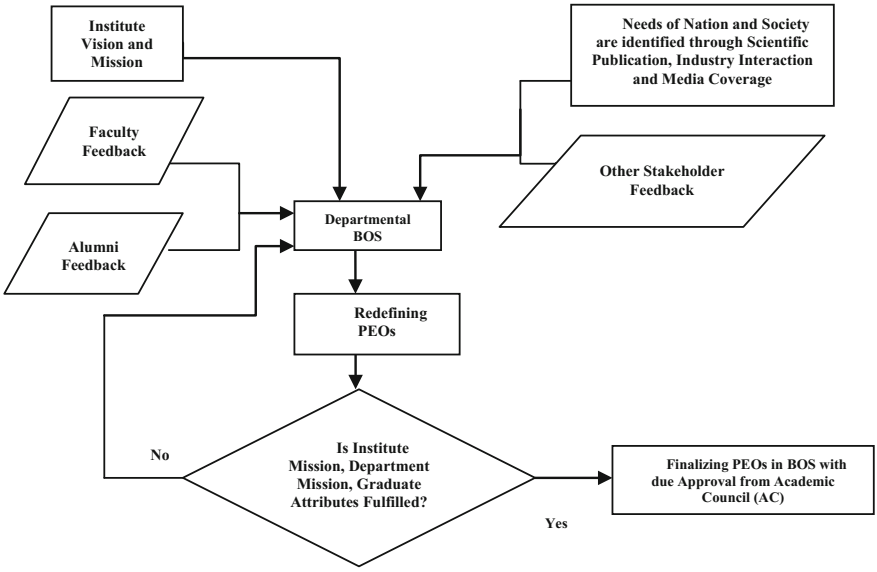


Fig. 2 Process employed for redefining the Program Educational Objectives

- **Proposed Process for Redefining the PEOs**

See Fig. 2.

- **Program Outcome (PO)**

Program Outcome must surge unswervingly and aid the Departmental mission. The connection amid the mission and the outcome should be unambiguous. Program Outcome must be directly allied to the academic discipline of the Program. Nevertheless, Program Outcome should be observable, measureable, and exclusively focussed on learning outcomes rather than curricular inputs.

- **Process Employed for Defining the Program Outcomes (PO)**

See Fig. 3.

- **Assessment Policy**
  - **Direct Assessment**
    1. **Observation of Student Knowledge/Student Evaluation**
       i. Continuous Evaluation Program
       ii. Class Test
       iii. Assignment

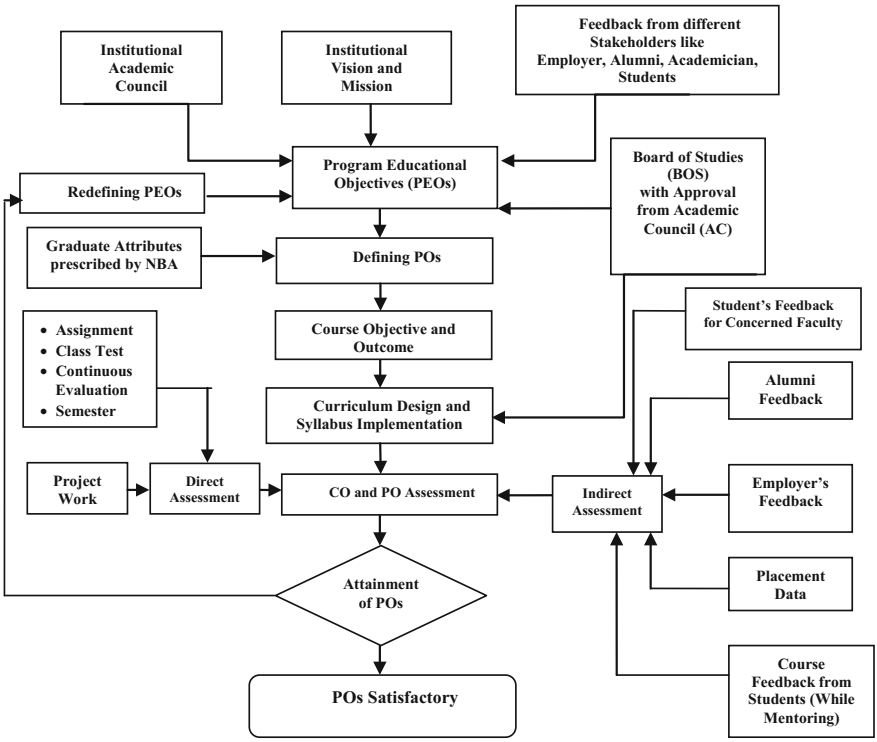


Fig. 3 Process employed for defining the Program Outcome

2. **Semester Examination**

– **Indirect Assessment**

1. **Student Feedback**

   i. Online Feedback Mechanism System exists for recording Feedback from Students for apiece Course and the Teacher in charge of the Course.
   ii. Feedback is recorded at the end of the Semester.
   iii. It is a parameter for gap analysis in the Program and attainment of POs.

2. **Alumni Feedback**

   i. Online Feedback Mechanism System exists for recording Feedback from Alumni for the Program.
   ii. Feedback is recorded in apiece year.
   iii. It is a parameter for gap analysis in the Program and attainment of POs.

3. **Employer Feedback**

   i. Online Feedback Mechanism System exists for recording Feedback from Employers about Students and Program.
   ii. Feedback is recorded in apiece year.
   iii. It is a parameter for gap analysis in the Program and attainment of POs.

4. **Faculty Feedback**

   i. Online Feedback Mechanism System exists for recording Feedback from Faculty Members for Individual Courses, thereby establishing attainment of Vision, Mission, PEOs, and POs.
   ii. Feedback is recorded in apiece year.

– **Process Employed for Designing the Program Curriculum**

Academic structure has been proposed according to the local requisites, keeping in wits the global academic standards. An assortment of feedback mechanism from diverse stakeholders has been amassed.

Board of Studies has been constituted comprising of faculty members, academic experts, and industry personnel. In BOS meetings, Feedback forms have been analyzed in presence of expert committee members consisting of academicians from leading colleges and universities along with industry personnel. The curriculum has been duly approved from Academic Council, for smooth functionality at Departmental level (Fig. 4).

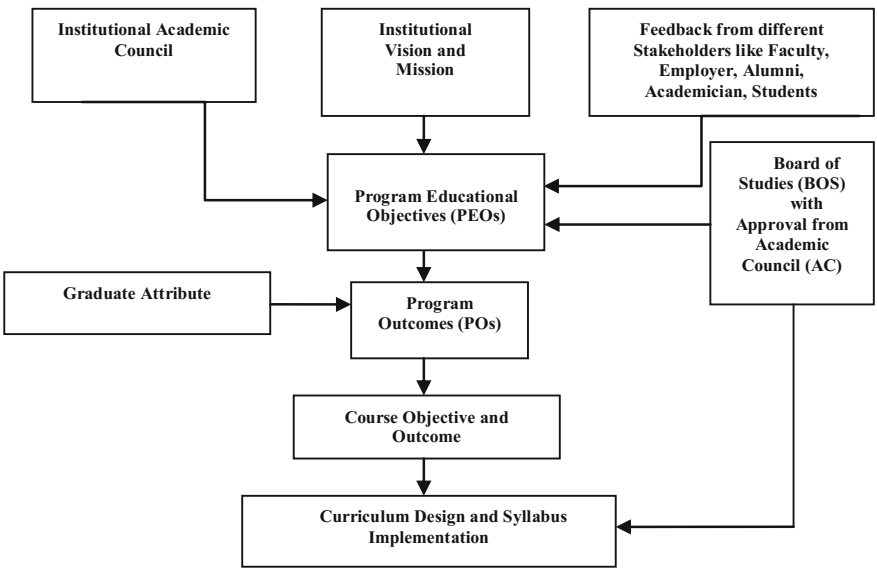


Fig. 4 Process employed for designing the Program Curriculum

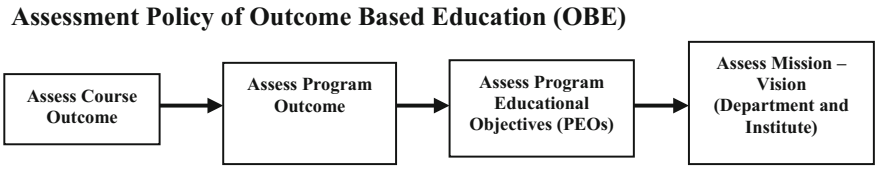


Fig. 5 Assessment Policy of Outcome-Based Education (OBE)

Assessment Policy of Outcome-Based Education (OBE)

See Fig. 5.

4 Conclusion

In this paper, realization of Outcome-Based Education (OBE) is ascertained as per requisite of learner fraternity. The syllabus might be customized according to the needs of learner and then initiated in the curriculum. The devised tailored curriculum should fulfill the entire vibrant requisite and thereby ensure the initiation of a new era in the field of education. Teaching and Learning methodology in engineering Institutions shape the spine of outcome and more precisely the realization of Outcome-Based Education (OBE). The inception of Outcome-Based Education (OBE) ensures proficiency in all aspects to craft the ambiance for the same. The conceptualization for the said work is initialized from a very fundamental notion that edification should not merely be imparted but should ascertain some sort of outcome, tagged with it. The outcome would signify the indispensable thrust for edification in its factual intellect.

References

1. Spady, W.: Leading and learning ‘outside the box’. Leadership in Focus, 4, Summer, pp. 23–25 (2006)
2. Spady, W.: Learning, leading and living outside the box: compelling insights from transformational research. In: Address to the Annual Conference of the Australian Primary Principal’s Association, Alice Springs, July 2006
3. Bransford, J., Brown, A., Cocking, R.: How People Learn. National Academy Press, Washington, D.C. (1999)
4. Pellegrino, J.W., Chudowsky, N., Glaser, R.: Knowing What Students Know: The Science and Design of Educational Assessments. National Academy Press, Washington, D.C. (2001)
5. Martin, J.: Needed: a paradigm for liberal education. In: Soltis, J.F. (ed.) Philosophy and Education, Yearbook of the National Society for the Study of Education, vol. 80, Part 1. National Society for the Study of Education, Chicago, IL (1981)
6. Manno, B.R.: Outcome Based Education: Miracle Cure or Plague? Hudson Institute Briefing Paper Number 165, June 1994

7. Australian Institute of Physics. Outcome Based Education (OBE) Philosophy a Disaster or Quality Education
8. Berlach, R.G.: Outcome based education and the death of knowledge. In: Annual Conference of the Australian Association for Research in Education Conference, The University of Melbourne, 28 Nov–2 Dec 2004
9. Louden, W., Chapman, E., Clarke, S., Cullity, M., House, H.: Evaluation of the Curriculum Improvement Program Phase 2. University of Western Australia, Crawley, WA (2007)
10. Bouslama, F., Lansari, A., Al–Rawi, A., Abonamah, A.: Assessing a new academic model using artificial neural networks. In: Proceedings of the 2002 IEEE International Conference on System, Man and Cybernetics, 6–9 Oct, Hammamet, Tunisia, WA1P6 (2002)

Author Index

A
Ahuja, Laxmi, 149
Archana, M., 357, 395
Athira, A.P., 367
Ayyasamy, A., 357, 395

B
Balamurali, S., 357
Ball, Amit Kumar, 477
Banerjee, Anuradha, 121
Banerjee, Meenakshi, 403
Banerjee, Sreeparna, 553
Basu, S., 89
Bera, Parthasarathi, 47
Bhattacharya, D.K., 561
Bhattacharya (Halder), Sharmistha, 209, 459
Bhattacharya, Ipshita, 135
Bhattacharya, Pinaki, 21
Bhattacharyya, Balaram, 497
Bhattacharyya, Sucharita, 229
Bhattacharyya, Swapan, 71
Bhuvaneswaran, Akshaya, 523
Biswas, Animesh, 283
Biswas, Aniruddha, 601
Biswas, Anuprava, 535
Biswas, Joyoti, 21
Biswas, Jyotirmoy, 47
Biswas, Papun, 59
Biswas, S., 173
Biswas, Swagata, 411
Bose, Rajesh, 309

C
Chakrabarti, Amlan, 185
Chakraborti, Debjani, 59
Chakraborty, Ritabrata, 81
Chanda, Reshmi, 59
Chatterjee, Punyasha, 411, 423
Chaudhuriin, Aditya, 449
Chhabra, Amit, 107
Choudhury, Nobhonil Roy, 561
Chowdhury, Rajdeep, 135, 609
Chowdhury, Ranjana, 21

D
Das, Antara, 403
Das, Bhaswati, 341
Das, G., 29
Das, Gourab, 11, 37
Das, Raju, 477
Das, Ria, 411
Das, Subhram, 561
Das, Sumit, 173, 601
Datta, Suprabeet, 323
De, M., 11, 29
De, Meenakshi, 37
De, Nirmita, 135
Debnath, Joy, 283
Debnath, Kalyani, 459
Devibala, K., 357
Dey, Abhishek, 511
Dey, Aritra, 173
Dey, Chanchal, 239
Dey, Kashi Nath, 511
Dey, Romi, 403
Deyasi, Arpan, 71, 81, 221, 341, 403
Dutta, Deepneha, 575
Dutta, Meghamala, 575
Dutta, Paramartha, 121
Dutta, Pratik, 163
Dutta, Shreya, 575
Dutta, Sourav, 575

G
Ganapathy, Sannasi, 485
Ganguly, Suchandan, 601
Ghosh, Prolay, 385
Ghosh, Saurav, 163

S. Bhattacharyya et al. (eds.), *Industry Interactive Innovations in Science, Engineering and Technology*, Lecture Notes in Networks and Systems 11,
DOI 10.1007/978-981-10-3953-9

Ghosh, Sayantani, 449
Ghosh, Sutanu, 299, 449
Ghoshal, Pranati, 435
Gopinath, Sakithya, 271
Goswami, Sanjay, 569
Goswami, Saptarsi, 185
Gupta, Rajashree, 591
Gupta, Somsubhra, 467, 545

H
Haldar, Sourish, 71
Halder, Debojyoti, 3

J
Jothi, Monika, 271

K
Kannan, Arputharaj, 485
Kathiriya, D.R., 349
Kaur, Bhajneet, 149
Kayal, Diptoneel, 553
Kishore Anthuvan Sahayaraj, K., 99
Kumar, Vinay, 149
Kumari, Barsha, 385

M
Maity, Goutam Kumar, 197
Majumder, Annwesha Banerjee, 467, 545
Malakar, Sourav, 185
Mandal, K.K., 11, 29, 37
Mandal, S., 11, 29, 37
Meikap, A.K., 89
Mitra, Amitava, 247
Mitra, Uddalak, 497
Mohan, R., 367
Mondal, Sanjoy, 163
Mudi, Rajani K., 239
Mukhopadhyay, Debapriya Sur, 59

N
Nag, Pooja, 255, 265
Naseera, Shaik, 523
Nath, Ujjwal Manikya, 239
Nath, Vijay, 377
Nayak, Dayananda, 265
Nayak, Shrabani, 81

O
Ohmprakash, V., 99
Oshin, 107

P
Pal, Utpal, 459
Panda, Ashis K., 247
Pandey, Shikha, 523
Paul, Aditi, 173
Paul, Arkadeep, 81
Paul, Pratiti, 221
Perumal, Varalakshmi, 271
Phadikar, Amit, 197
Phadikar, Baisakhi Sur, 197
Prajapati, Nilesh B., 349
Prasad, Deepak, 377
Premkumar, M., 247

R
Raha, Bipasa, 3
Rajini, G.K., 523
Ramachandran, Nandhakumar, 271
Roy, Rajat K., 247
Roy, Shibendu Shekhar, 477
Roy, Srijita Barman, 209

S
Sadani, Kapil, 255, 265
Saha, Soumyabrata, 609
Saha, Subhajit, 135
Saha, Tanusree, 385
Sahana, Sudipta, 309, 609
Sajid Sarwar, Mir Md., 423
Sangeetha, Ganesan, 485
Sanki, Debashis, 609
Sanyal, Avik, 591
Sarddar, Debabrata, 309
Sarkar, Indrani, 569
Sathish, S., 395
Sen, Sunit Kumar, 435
Shinde, Sumit, 255, 265
Sikdar, Swati, 575
Sinha, R., 89
Sufian, Abu, 121

T
Thander, Anup Kumar, 229
Tibarewala, D.N., 561
Tudu, B., 29

V
Venkatachalapathy, K., 99
Verma, Pratibha, 221
Vijayalakshmi, Muthuswamy, 485
Vijayan, Midhula, 367