

Durbadal Mandal

Rajib Kar

Swagatam Das

Bijaya Ketan Panigrahi *Editors*

Intelligent Computing and Applications

Proceedings of the International
Conference on ICA, 22–24 December
2014



Springer

Advances in Intelligent Systems and Computing

Volume 343

Series editor

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

About this Series

The series “Advances in Intelligent Systems and Computing” contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing.

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

Advisory Board

Chairman

Nikhil R. Pal, Indian Statistical Institute, Kolkata, India
e-mail: nikhil@isical.ac.in

Members

Rafael Bello, Universidad Central “Marta Abreu” de Las Villas, Santa Clara, Cuba
e-mail: rbellop@uclv.edu.cu

Emilio S. Corchado, University of Salamanca, Salamanca, Spain
e-mail: escorchedo@usal.es

Hani Hagras, University of Essex, Colchester, UK
e-mail: hani@essex.ac.uk

László T. Kóczy, Széchenyi István University, Győr, Hungary
e-mail: koczy@sze.hu

Vladik Kreinovich, University of Texas at El Paso, El Paso, USA
e-mail: vladik@utep.edu

Chin-Teng Lin, National Chiao Tung University, Hsinchu, Taiwan
e-mail: ctlin@mail.nctu.edu.tw

Jie Lu, University of Technology, Sydney, Australia
e-mail: Jie.Lu@uts.edu.au

Patricia Melin, Tijuana Institute of Technology, Tijuana, Mexico
e-mail: epmelin@hafsamx.org

Nadia Nedjah, State University of Rio de Janeiro, Rio de Janeiro, Brazil
e-mail: nadia@eng.uerj.br

Ngoc Thanh Nguyen, Wroclaw University of Technology, Wroclaw, Poland
e-mail: Ngoc-Thanh.Nguyen@pwr.edu.pl

Jun Wang, The Chinese University of Hong Kong, Shatin, Hong Kong
e-mail: jwang@mae.cuhk.edu.hk

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

Durbadal Mandal · Rajib Kar
Swagatam Das · Bijaya Ketan Panigrahi
Editors

Intelligent Computing and Applications

Proceedings of the International Conference
on ICA, 22–24 December 2014



Springer

Editors

Durbadal Mandal
Department of Electronics and
Communication Engineering
NIT
Durgapur, West Bengal
India

Rajib Kar
Department of Electronics and
Communication Engineering
NIT
Durgapur, West Bengal
India

Swagatam Das
Machine Intelligence Unit
ISI
Kolkata, West Bengal
India

Bijaya Ketan Panigrahi
Department of Electrical Engineering
IIT Delhi
New Delhi
India

ISSN 2194-5357

ISSN 2194-5365 (electronic)

Advances in Intelligent Systems and Computing

ISBN 978-81-322-2267-5

ISBN 978-81-322-2268-2 (eBook)

DOI 10.1007/978-81-322-2268-2

Library of Congress Control Number: 2015930027

Springer New Delhi Heidelberg New York Dordrecht London
© Springer India 2015

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

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

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

Printed on acid-free paper

Springer (India) Pvt. Ltd. is part of Springer Science+Business Media (www.springer.com)

Preface

This AISC volume contains the papers presented at the First International Conference on Intelligent Computing and Application (ICICA 2014) held during December 22–24, 2014, at NIT Durgapur, West Bengal, India. ICICA 2014 is the first international conference aiming at bringing together researchers from academia and industry to report and review the latest progresses in the cutting-edge research on various research areas of Image, Language Processing, Computer Vision and Pattern Recognition, Machine Learning, Data Mining and Computational Life Sciences, Management of Data including Big Data and Analytics, Distributed and Mobile Systems including Grid and Cloud infrastructure, Information Security and Privacy, VLSI, Electronic Circuits, Power Systems, Antenna, Computational fluid dynamics and Heat transfer, Intelligent Manufacturing, Signal Processing, Intelligent Computing, Soft Computing, Bio-informatics, Bio Computing, Web Security, Privacy and E-Commerce, E-governance, Service Orient Architecture, Data Engineering, Open Systems, Optimization, Communications, Smart wireless and sensor Networks, Smart Antennae, Networking and Information security, Machine Learning, Mobile Computing and Applications, Industrial Automation and MES, Cloud Computing, Green IT, IT for Rural Engineering, Business Computing, Business Intelligence, ICT for Education for solving hard problems, and finally to create awareness about these domains to a wider audience of practitioners.

ICICA 2014 received 144 paper submissions including two foreign countries across the globe. After a rigorous peer-review process involving 600 reviews, 70 full length articles were accepted for oral presentation at the conference. This corresponds to an acceptance rate of 48.61 % and is intended to maintain the high standards of the conference proceedings. The papers included in this AISC volume cover a wide range of topics in intelligent computing and algorithms and their real-world applications in problems from diverse domains of science and engineering.

The conference featured distinguished keynote speakers: Dr. P.N. Suganthan, NTU, Singapore; Dr. Swagatam Das, ISI, Kolkata, India; Prof. B.K. Panigrahi, IIT Delhi, India and Prof. B. Sikdar, Rensselaer Polytechnique Institute, USA. We take this opportunity to thank the authors of the submitted papers for their hard work, adherence to the deadlines, and patience with the review process.

The quality of a referred volume depends mainly on the expertise and dedication of the reviewers. We are indebted to the Program Committee/Technical Committee members, who produced excellent reviews in short time frames. We would also like to thank TEQIP II for providing all the logistical support and financial assistance. First, we are indebted to NIT Durgapur, India for supporting our cause and encouraging us to organize the conference there. In particular, we would like to express our heartfelt thanks for providing us with the necessary financial support and infrastructural assistance to hold the conference. Our sincere thanks to Prof. P. Gupta, Dean Research and Consultancy; Prof. D.K. Mandal; Prof. J.P. Sarkar; Prof. B. Halder; Col. (Retd.) P.S. Sandhu (Registrar); Prof. A.K. Bhattacharjee; Prof. R. Ghatak; Prof. K.C. Ghanta; Prof. N.K. Roy; Dr. (Mrs.) S. Roy; Prof. S. Ghosh; Mr. A.K. Chattopadhyay; Mr. U.C. Mukherjee; Mr. A. Kumar; Mr. A.K. Bhagat and Prof. T. Kumar (Director, NIT Durgapur), for their encouragement and continuous support. We thank Maurice Clerc, France; Roderich Gross, England; Kalyanmoy Deb, India; Saman Halgamuge, Australia; Jeng-Shyang Pan, Taiwan; Peng Shi, UK; Javier Del Ser, Spain; Leandro Dos Santos Coelho, Brazil; S.S. Pattanaik, India; Gerardo Beni, USA; K. Parsopoulos, Greece; Lingfeng Wang, China; Athanasios V. Vasilakos, Athens; Pei-Chann Chang, Taiwan; Chilukuri K. Mohan, USA; Saeid Nahavandi, Australia; Abbas Khosravi, Australia; Almoataz Youssef Abdelaziz, Egypt; K.T. Chaturvedi, India; M.K. Tiwari, India; Yuhui Shi, China; A. Anand, India; Dipankar Dasgupta, USA; Lakhmi Jain, Australia; X.Z. Gao, Finland; Juan Luis Fernandez Martinez, Spain; Oscar Castillo, Mexico; Heitor Silverio Lopes, Brazil; S.K. Udgata, India; Namrata Khemka, USA; G.K. Venayagamoorthy, USA; Zong Woo Geem, USA; Ying Tan, China; S.G. Ponnambalam, Malaysia and Halina Kwasnicka, Poland for providing valuable guidelines and inspiration to overcome various difficulties in the process of organizing this conference.

We would also like to thank the participants of this conference. Finally, we thank all the volunteers who made great efforts in meeting the deadlines and arranging every detail to make sure that the conference could run smoothly. We hope the readers of these proceedings find the papers inspiring and enjoyable.

December 2014

Durbadal Mandal
Rajib Kar
Swagatam Das
Bijaya Ketan Panigrahi
P.N. Suganthan

Conference Organizing Committee

Chief Patron

Prof. Bikas Sinha, Chairman, NIT Durgapur, India

Patron

Prof. T. Kumar, Director, NIT Durgapur, India

General Chair

Dr. D. Mandal, Dr. R. Kar, NIT Durgapur, India

General Co-chair

Prof. S. Das, ISI Kolkata, India

Prof. B.K. Panigrahi, IIT Delhi, India

Program Chair

Prof. D.K. Pratihar, IIT Kharagpur, India
Prof. Jaya Sil, IIEST Shibpur, India
Prof. S. Kundu, NIT Durgapur, India

Finance Chair

Dr. Rajat Mahapatra, NIT Durgapur, India

Publication Chairs

Prof. S.P. Ghoshal, Dr. A.K. Mal, NIT Durgapur, India

International Advisory Committee/Technical Committee

Maurice Clerc, France
Roderich Gross, England
Kalyanmoy Deb, India
Saman Halgamuge, Australia
Jeng-Shyang Pan, Taiwan
Peng Shi, UK
Javier Del Ser, Spain
Leandro Dos Santos Coelho, Brazil
S.S. Pattanaik, India
Gerardo Beni, USA
K. Parsopoulos, Greece
Lingfeng Wang, China
Athanasios V. Vasilakos, Athens
Pei-Chann Chang, Taiwan
Chilukuri K. Mohan, USA
Saeid Nahavandi, Australia
Abbas Khosravi, Australia
Almoataz Youssef Abdelaziz, Egypt
K.T. Chaturvedi, India
M.K. Tiwari, India
Yuhui Shi, China
A. Anand, India

Dipankar Dasgupta, USA
Lakhmi Jain, Australia
X.Z. Gao, Finland
Juan Luis Fernandez Martinez, Spain
Oscar Castillo, Mexico
Heitor Silverio Lopes, Brazil
S.K. Udgata, India
Namrata Khemka, USA
G.K. Venayagamoorthy, USA
Zong Woo Geem, USA
Ying Tan, China
S.G. Ponnambalam, Malaysia
Halina Kwasnicka, Poland

Organizing Committee

Prof. P. Gupta, NIT Durgapur
Prof. D.K. Mandal, NIT Durgapur
Prof. J.P. Sarkar, NIT Durgapur
Prof. B. Haldar, NIT Durgapur
Col. (Retd.) P.S. Sandhu, NIT Durgapur
Prof. A.K. Bhattacharjee, NIT Durgapur
Prof. R. Ghatak, NIT Durgapur
Prof. K.C. Ghanta, NIT Durgapur
Prof. N.K. Roy, NIT Durgapur
Dr. (Mrs.) S. Roy, NIT Durgapur
Prof. S. Ghosh, NIT Durgapur
Mr. A.K. Chattopadhyay, NIT Durgapur
Mr. U.C. Mukherjee, NIT Durgapur
Mr. A. Kumar, NIT Durgapur
Mr. A.K. Bhagat, NIT Durgapur

Contents

Characterization of Fuzzy Tree Searches: A Perspective Note	1
Parag Bhalchandra, Santosh Khamitkar, Nilesh Deshmukh, Sakharam Lokhande and Satish Mekewad	
A Modified SOM-Based RBFN for Rotation Invariant Clear and Occluded Fingerprint Recognition	11
Sumana Kundu and Goutam Sarker	
An Unsupervised OCA-based RBFN for Clear and Occluded Face Identification	19
Dhananjay Bhakta and Goutam Sarker	
Analysing the Impact of Human Behaviour over Machine Learning	27
Shailendra Singh Yadav and Bhupendra Verma	
An Improved RC4 with Statistical Analysis on Ciphertexts	37
Suman Das, Hemanta Dey and Ranjan Ghosh	
Gathering of Swarm of Mobile Robots in Presence of Horizontal Obstacles	49
U. Baveenther and Deepanwita Das	
Design of Wireless Sensor Node to Measure Vibration and Environment Parameter for Structural Health Monitoring Application	59
Sen Siddheswar, Swarnakar Biplab and Datta Uma	
Economic Load Dispatch Considering Non-smooth Cost Functions Using Predator-Prey Optimization	67
Sunanda Hazra and Provas Kumar Roy	

Sparse Denoising in Speckle Noise: A Comparative Study of Dictionaries	79
Suchismita Maiti, Amish Kumar and Debasish Nandi	
Compact RS(32, 28) Encoder	89
Jagannath Samanta, Jaydeb Bhaumik and Soma Barman	
Graphene–Silicene Composite can Increase the Efficiency of Cloud Computing	97
Dhiman Kumar Das and Sukadev Sahoo	
Object-oriented Modeling of IDEA for E-learning Security	105
Ambalika Ghosh and Sunil Karforma	
An Improved BAT-Optimized Cluster-Based Routing for Wireless Sensor Networks	115
Koteswararao Seelam, M. Sailaja and T. Madhu	
An Hybrid Ant Routing Algorithm for Reliable Throughput Using MANET	127
N. Umapathi, N. Ramaraj, D. Balasubramaniam and R. Adlin mano	
Development of Type 2 Fuzzy Rough Ontology-based Middleware for Context Processing in Ambient Smart Environment	137
A.B. Karthick Anand Babu and R. Sivakumar	
Multiple-Image Encryption Using Genetic Algorithm	145
Subhajit Das, SatyendraNath Mandal and Nabin Ghoshal	
Knowledge-Based System Architecture on CBR for Detection of Cholera Disease	155
Souvik Chakraborty, Chiranjit Pal, Shombo Chatterjee, Baisakhi Chakraborty and Nabin Ghoshal	
Efficient Adaptive Routing Algorithm for WSNs with Sink Mobility	167
Ashwini V. Mannapur, Jayashree D. Mallapur and S.P. Parande	
Detection of Degree of Sickness of Affected Eye Using Fuzzy Logic and Histogram Analysis	177
Manisha Barman, J. Paul Choudhury and S. Biswas	

Element Spacing Optimization of Low Sidelobe Concentric Hexagonal Antenna Arrays Using MOEA/D	189
Sudipta Das, Durbadal Mandal, Rajib Kar and Sakti Prasad Ghoshal	
QPSO for Synthesis of Scanned Linear Array Antenna for Fixed Side Lobe Level and First Null Beam Width Including Wide Null Placement	197
Hemant Patidar and Gautam Kumar Mahanti	
Decision Feedback Equalization for MIMO Systems	205
Samarendra Nath Sur, Rabindranath Bera and Bansibadan Maji	
Fuzzy Logic-Based Handover in 3GPP LTE Network	213
Parul Datta and Sakshi Kaushal	
FPGA Implementation of Novel Discrete Phase-Locked Loop	223
N. Bharani dharan, M. Chinnathambi and S. Rajaram	
Analytical Approach on the Scale Length Model for Tri-material Surrounding Gate Tunnel Field-Effect Transistors (TMSG-TFETs)	231
P. Vanitha, G. Lakshmi Priya, N.B. Balamurugan, S. Theodore Chandra and S. Manikandan	
Color Image Segmentation Using Cauchy-Mutated PSO	239
Suman Banerjee, Debasree Saha and Nanda Dulal Jana	
Large-Scale Global Optimization Using Dynamic Population-Based DE	251
Seema Chauhan, Suman Banerjee and Nanda Dulal Jana	
A Novel Preprocessing Approach for Human Face Recognition Invariant to Illumination	261
U.K. Jaliya and J.M. Rathod	
Modeling and Simulation of Semi-active Jerk-Driven Damper on Severe Speed Breaker	267
C.K. Nirala and S. Mandal	
Voltage-Controlled Ring Oscillator for Harmonic Frequency Generation	277
Surajit Mal, Ashis Kumar Mal and Sumalya Ghosh	

Amplifier Design Optimization in CMOS	287
Sumalya Ghosh, Ashis Kumar Mal and Surajit Mal	
An Application of ANFIS-Based Intelligence Technique for Predicting Tool Wear in Milling	299
Shibendu Shekhar Roy	
A High-performance Elliptic Curve Cryptographic Processor for FPGA Platform	307
V.R. Venkatasubramani, N. Premkumar, K. Vignesh and S. Rajaram	
Multi-focus and Multi-exposure based Color Image Fusion for Concealed Weapon Detection	313
Ekta M. Upadhyay and N.K. Rana	
A Compact Analytical Model for 2D Triple Material Surrounding Gate Nanowire Tunnel Field Effect Transistors	325
D. Saraswathi, N.B. Balamurugan, G. Lakshmi Priya and S. Manikandan	
Application of Particle Swarm Optimization Technique in Hexagonal and Concentric Hexagonal Antenna Array for Side Lobe Level Reduction	333
Rajesh Bera, Durbadal Mandal, Rajib Kar and Sakti Prasad Ghoshal	
Bearing-Only Tracking Using Sparse-Grid Gauss–Hermite Filter	349
Rahul Radhakrishnan, Shovan Bhaumik, Nutan Kumar Tomar and Abhinoy Kumar Singh	
Novel Design and Implementation of Passive Infrared Sensor in Steel Industry Automation	357
Basava Naga Girish Koneru, Vijaya Krishna Teja Mantripragada, Prithvi Raj Pani and G. Lakshminarayanan	
A Comparative Study of Feature Selection and Machine Learning Methods for Sentiment Classification on Movie Data Set	367
C. Selvi, Chakshu Ahuja and E. Sivasankar	
Enhancement of Transmission System Loadability During Contingency by Optimal Allocation of FACTS Devices Using Particle Swarm Optimization	381
P. Malathy, A. Shunmugalatha and P. Thaineesh	

Fault Diagnosis of Broken Rotor Bars in Induction Motor Using Multiscale Entropy and Backpropagation Neural Network	393
Alok Verma and Somnath Sarangi	
Design of a Signal Sensor for Analyzing Biological Activities at Cellular Level	405
Amit Krishna Dwivedi, Anubhav Sinha and Aminul Islam	
Fuzzy PID Control for Ball and Beam Mechanism	413
Nirmal Baran Hui and Pratima Sharma	
Framework for Smart Health: Toward Connected Data from Big Data	423
P. Vignesh Raja, E. Sivasankar and R. Pitchiah	
Capacity Analysis of ZF Receiver in Correlated MIMO Nakagami-m Channel	435
Samarendra Nath Sur, Soumyasree Bera, Rabindranath Bera and Bansibadan Maji	
Neural Estimations of First Null Beamwidth for Broadside and End-Fire Uniform Linear Antenna Arrays	445
Subhash Mishra, Ram Narayan Yadav and Rajendra Prasad Singh	
Representing the Extended H_∞ Filter (EHF) as a Modified Extended Kalman Filter (EKF) for Predicting the Range of the Robustness Bound	453
Manika Saha, Ratna Ghosh and Bhaswati Goswami	
Performance Evaluation of Forward Error Correction Schemes in Wireless Body Area Networks	469
K.S. Deepak and A.V. Babu	
Simulating Self-Recovering Electric Circuits Using Neural Networks	479
Arijit Ray, S.S. Dash and N. Chellammal	
Dynamic Scenarios in Embedded System Verification	493
Hara Gopal Mani Pakala, P.L.H. Varaprasad, K.V.S.V.N. Raju and Ibrahim Khan	

PSO in Concentric Circular Arrays for Side Lobe Reduction with Symmetric Relocation Boundary Condition	503
Pragnan Chakravorty and Durbadal Mandal	
Ear Recognition Using Force Field Transform and Collaborative Representation-Based Classification with Single Training Sample Per Class	511
Sayan Banerjee and Amitava Chatterjee	
Demonstration of GPGPU-Accelerated Computational Fluid Dynamic Calculations	519
Samarth Agrawal, Manish Kumar and Somnath Roy	
High-performance Current Mode Receiver Design for On-chip VLSI Interconnects	527
Yash Agrawal, Rajeevan Chandel and Rohit Dhiman	
CMOS Amplifier Design Using Simplified g_m/I_D Technique	537
Agnish Mal, Ashis Kumar Mal and Sanjit Kumar Datta	
Solution of Optimal Power Flow by an Opposition-Based Gravitational Search Algorithm	545
Binod Shaw, Abhik Banerjee, V. Mukherjee and S.P. Ghoshal	
Information Systems and Artificial Intelligence Technology Applied in Concrete Road Design	559
Erhan Burak Pancar and Muhammet Vefa Akpinar	
Multi-hop Secured Video Transmission Using Dual-channel Dual-path Routing Against Packet Copy Attack	569
Vaibhav Mishra, Saurabh Maheshwari and Kavita Choudhary	
Static Analysis: A Survey of Techniques and Tools	581
Anjana Gosain and Ganga Sharma	
A LabVIEW-Based Data Acquisition System in a Quarter Car Test Rig to Optimize Vehicle Suspension System	593
Anirban C. Mitra and Nilotpal Banerjee	
A Computational Structural Biology of SoxR and DNA: A Modelling and Interactive Discern to Express the Sox Operon in <i>Pseudaminobacter salicylatoxidans</i> (KCT001) for Global Sulphur Oxidation	603
Sujay Ray, Arundhati Banerjee and Angshuman Bagchi	

Contents	xvii
Empirical Evaluations Using Character and Word N-Grams on Authorship Attribution for Telugu Text.	613
S. Nagaprasad, T. Raghunadha Reddy, P. Vijayopal Reddy, A. Vinaya Babu and B. VishnuVardhan	
Author Index	625

About the Editors

Durbadal Mandal is Assistant Professor in the Department of Electronics and Communication, National Institute of Technology, Durgapur, India. He completed his B.E. in Electronics and Communication Engineering, from Regional Engineering College, Durgapur in 1996. He received the M.Tech. and Ph.D. degrees from National Institute of Technology, Durgapur, in 2008 and 2011, respectively. His research interest areas include array antenna design and filter optimization via evolutionary computing techniques. He has over 240 research publications in Journals and Conferences of international repute.

Rajib Kar received his B.E. degree in Electronics and Communication Engineering, from Regional Engineering College, Durgapur, West Bengal, India in the year 2001. He received the M.Tech. and Ph.D. degrees from National Institute of Technology, Durgapur, West Bengal, India in the year 2008 and 2011, respectively. Presently, he is working as Assistant Professor of electronics and communication engineering at National Institute of Technology, Durgapur. His research interest area include VLSI signal processing, filter optimization via evolutionary computing techniques. He has supervised two Ph.D. students. He has published over 250 research papers in International Journals and Conferences.

Dr. Swagatam Das received the B.E. Tel. E., M.E. Tel. E (Control Engineering specialization) and Ph.D. degrees, all from Jadavpur University, India, in 2003, 2005, and 2009, respectively. Currently, he is serving as Assistant Professor at the Electronics and Communication Sciences Unit of Indian Statistical Institute, Kolkata. His research interests include evolutionary computing, pattern recognition, multi-agent systems, and wireless communication. Dr. Das has published one research monograph, one edited volume, and more than 150 research articles in peer-reviewed journals and international conferences. He is the founding co-editor-in-chief of “Swarm and Evolutionary Computation”, an international journal from Elsevier. He serves as associate editor of the IEEE Trans. on Systems, Man, and Cybernetics: Systems and Information Sciences (Elsevier). He is an editorial board member of Progress in Artificial Intelligence (Springer), Mathematical Problems in Engineering, International Journal of Artificial Intelligence and Soft Computing,

and International Journal of Adaptive and Autonomous Communication Systems. He is the recipient of the 2012 Young Engineer Award from the Indian National Academy of Engineering (INAE).

Dr. Bijaya Ketan Panigrahi is Associate Professor of Electrical and Electronics Engineering Department in Indian Institute of Technology Delhi, India. He received his Ph.D. degree from Sambalpur University. He is serving as chief editor in the International Journal of Power and Energy Conversion. His interests focus on power quality, FACTS devices, power system protection, and AI application to power system.

Characterization of Fuzzy Tree Searches: A Perspective Note

Parag Bhalchandra, Santosh Khamitkar, Nilesh Deshmukh,
Sakharam Lokhande and Satish Mekewad

Abstract The objective of this paper was to deal with the imprecise data involved in different kinds of searching techniques in a more efficient way and thus to characterize an innovative version of search under the uncertainty. This paper is one of the outcomes of a rigorously reviewed experimental work carried out on the synthesis of constraint satisfaction problems through tree searching algorithms where it was found that the traditional tree search algorithms fail to search in uncertain problem domain.

Keywords Tree searching · Complexity · Decision trees · Fuzzy trees

1 Introduction

The real-world data analysis including data mining and classification problems usually involves a large number of candidate inputs or features, which are almost nonlinear and are difficult enough to model [1]. Besides nonlinearity, sometimes it is difficult to choose the most relevant input variables. Thus, input selection is a crucial step with the aim of reducing the model's complexity and removing inputs

P. Bhalchandra (✉) · S. Khamitkar · N. Deshmukh · S. Lokhande · S. Mekewad
School of Computational Sciences, S.R.T.M. University, Nanded 431606, Maharashtra, India
e-mail: srtmun.parag@gmail.com

S. Khamitkar
e-mail: s.khamitkar@gmail.com

N. Deshmukh
e-mail: nileshkd@yahoo.com

S. Lokhande
e-mail: sana_lokahande@rediffmail.com

S. Mekewad
e-mail: satishmekewad@gmail.com

that do not contribute to the model performance. Input selection methods for nonlinear systems have been studied in many places [1]. These methods can generally be divided in two main groups: model-free methods and model-based methods [2]. Usually, these methods are straightforward and allow us to test all the subsets of variables. While testing, we observe that there are two different approaches, namely top-down tree search and bottom-up tree search. Both methods perform input selection on applications such as classification and nonlinear dynamic modeling. The model obtained at every iteration is constructed using uncertainty modeling, where the rules are identified initially. In today's era, acquisition of knowledge is a big issue that can be improved by innovating good data structures so as to reduce the search complexity of traditional way of searching for acquisition of knowledge [2]. Approaches such as symbolic decision trees for decision making and classification exist today with their inference procedures. This is very useful for people who are interested in knowledge domain understanding, classification, or the symbolic rules. This leads to extensive research efforts resulting in a number of methodological and empirical advancements [3]. Decision trees were proposed by Quinlan [4–6] with the ID3 program. These use symbolic domains and hence are not useful for numeric decisions.

2 Review of Existing Approaches

In light of above discussions, we feel that the traditional trees need to be modified so as to reduce the tree search complexity values. This is possible only if we insist for intelligent data analysis. To do so, we need to modify existing terminology and structure of trees. This results in new types of trees such as decision trees, fuzzy trees, and fuzzy decision trees [1, 7]. For efficient searching in trees, we need to modify the structural ambience and propose handling of uncertainties. This can be accomplished by using decision trees as models and fuzzy mechanism as navigation mechanism. Many areas such as subset selection, classification, and regression analysis make use of decision trees. These trees use graphical model and text rules for navigation [7]. Since they are easy to understand and navigate, they require less computational resources as compared to traditional ones. Fuzzy systems are based on fuzzy logic and fuzzy sets, which allow us to deal with uncertain and fuzzy data. The fuzzy decision models are hybrid of fuzzy systems and decision trees. These are suitable for providing simple, accurate, and highly interpretable models at low computational costs [7]. As proposed by Prof. Zadeh [8], a fuzzy system has at least one of its attributes defined by fuzzy logic, fuzzy set theory, and fuzzy sets. A fuzzy system is usually composed of a knowledge base and an inference mechanism [8]. The knowledge base contains a fuzzy rule base and a fuzzy data base [8]. The former has the rules that form the core of the system [8, 9]. These rules are constructed based on the fuzzy sets defining the attributes of the system, stored in the fuzzy data base which along with inference mechanism helps to classify new examples. Fuzzy decision trees have also been proposed in the literature [3, 9, 10].

There exist many algorithms to construct decision trees. One of the most used algorithms for constructing decision trees has long been the ID3 method introduced by Quinlan [4]. This algorithm tries to construct the smallest classification tree based on a small set of training examples. The main disadvantage is that this algorithm only considers symbolic discrete values both for attributes and for the resulting class. Another major disadvantage of this method includes the fact that the pre-partitioning requires previous knowledge of the data, that is, how many intervals are required and how to choose the split points between the intervals. Subsequent improvements of the ID3 algorithm, such as C4.5 [7] and CART [1, 9], try to deal with intervals in an optimal manner as they always try to split a compact interval into two subintervals, and they accomplish this by choosing the optimal split point. In terms of interpretability and ability of processing uncertainty and imprecision of fuzzy systems, fuzzy trees have been evolved as powerful models for handling uncertainty [9]. Moreover, fuzzy decision trees inherit the desirable characteristics of decision trees regarding their low computational induction cost, as well as the possibility of expressing the induced models graphically and as a set of rules [3, 8, 11].

3 Understanding Fuzzification Process

Traditionally, decision trees have two major aspects, that is, a procedure to build the symbolic tree and a method for decision making. We will take into consideration some of the approaches catering to both components. The neural networks can be an alternative to our approach due to ease of application and ability to provide responses. However, their clarity level does not match [7]. This might be a problem when decisions need to be justified. For such situations, we need to innovate a new approach to combine some elements of symbolic and subsymbolic approaches. In decision trees, the resulting tree can be cut or pruned or restructured, which often leads to improved generalization by incorporating additional bias [11]. In fuzzy representation, knowledge can be optimized at various levels: fuzzy set definitions, norms used, etc. [2, 11]. In order to understand these concepts, we need to understand the basic constituents first, which are attempted in below sections. The concept of domain in traditional machine learning systems and in fuzzy systems is totally different. The domain for attributes can be discrete in traditional systems, whereas they are a range or continuous in fuzzy systems. Some algorithms work for discrete values and some work for continuous ones [2, 11]. Interchanging a continuous domain with a discrete one or vice versa is not very easy and needs a good understanding of the knowledge domain. This is called as covering of domain, and it is usually carried out by partitioning and clustering methods [2, 11]. While working for conversion process, the covering must put every domain element into one of the covered subset. If so, the covering is called as complete. This holds true in traditional approach but false in fuzzy sets. If the covering is inconsistent, then a domain value can be found in more than one subset. This holds true in fuzzy sets.

Once fuzzy comes in picture, we derive the subset/domains. We then label each derived subsets. In general, a discrete domain can be unordered, ordered partially, or ordered completely as will every discretized domain [2, 9]. Any ordering can be advantageous as it can be used to define the notion of similarity. For example, consider the attribute skin color [11], with discrete values: white, tan, and black. In this case, one may say that white is “more similar” to tan than to black. Such orderings, along with similarity measures, are being used to advantage in many algorithms. They will also be used to advantage in fuzzy trees (Fig. 1).

To deal with such ambiguities, additional numeric components have been added in symbolic systems [3, 11]. This is the reason due to which decision trees have been extended to accommodate probabilistic measures [12] and decision rules have been extended by flexible matching mechanisms [11].

As another alternative, we can think of fuzzy sets. The universe of disclosure [11] represents the degree to which a set belongs. This is also known as fuzzy subset. The labels for fuzzy subsets are described by fuzzy linguistic variables, which are essentially attributes called as fuzzy terms. Their domain contains linguistic values [11]. We can take a fuzzy term alternatively as a fuzzy set, as the fuzzy set itself is defined by its membership function [3, 10]. For example [11], consider an attribute income. It can be continuous or discrete attribute. If we use three linguistic terms such as low, medium, and high, it becomes fuzzy. The low, medium, and high are called as domain values. The fuzzy sets being defined similarly overlap to reflect fuzziness of the concepts and should cover the whole universe of discourse for completeness [11]. This is illustrated in Fig. 2.

Figure 2 above shows the possibilities for actual income. As understood, the actual income can be low, medium, and high. But the income cannot be discrete all the

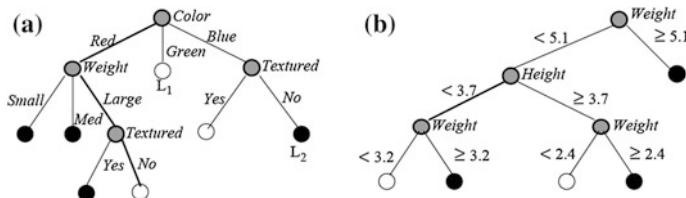


Fig. 1 Understanding fuzzy trees for the skin color example

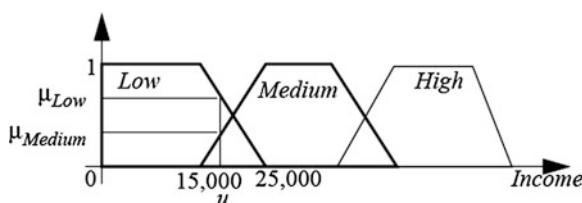


Fig. 2 Fuzzification of the income

times. There can be fuzzification actual income as low and medium both with different degrees. Fuzzy sets are generally described with convex functions [11] peaking at 1 (normal fuzzy sets). Trapezoidal (or triangular) sets are most widely used [11], in which case a normal set can be uniquely described by the four (three) corners values of U. For example, low in Fig. 2 could be represented as $<0,0,10,000,25,000>$.

In computer storage affiance, trapezoidal sets are very popular [7, 11]. Fuzzy sets are capable of defining classical set operations including union, intersection, and complements. The characteristic functions can be used for defining such operations. However, due to the infinite number of membership values, infinitely many interpretations can be assigned to those operations [11]. Fuzzy propositions work similar to classical propositions. We can combine them with logical connectives to form fuzzy statements. Two fuzzy sets intersect with two linguistic values if both of the fuzzy variables share the same universe of discourse. In other situation, the relationship expressed in the statement can be represented with a fuzzy relation [11]. The remarkable observation about fuzzy relations can be made that fuzzy sets satisfying the relation can be induced from fuzzy sets satisfying other elements of the relation by means of composition [7, 11, 12], provided that the relation is defined at least once. This helps us to express the knowledge with similar rules [11].

We will address such interpretation in decision and fuzzy trees. We will stick to the narration made in contemporary works [13]. A simple way to explain how is the characterization of the rules that are inferred from a decision tree is to consider all the leaves of the tree. For each leaf, a conjunction can be easily constructed by reading the labels of the nodes and branches that are encountered starting from the root of the tree down to the corresponding leaf. Then, in order to get the condition for one of the classes, we would normally construct a disjunction between all the leaves that have a value associated with that class. One of the major disadvantages of classical decision trees is that each tested object will have only one leaf associated with it. Moreover, this leaf is closely related to the input samples. In many cases, it would be a nice feature to have close values somehow related to each other. Also, it would be nice to be able to construct decision trees by using symbolic values such as tall and short, and hot and cold. All these issues are solvable using a fuzzy approach [4, 5]. Going back to the classical decisions inferred by a decision tree, in order to test an object against a given decision tree, we would start from the root of the tree and go down the tree by testing the attribute corresponding to the current node and following only the branch of the tree that corresponds to the value that has the same value as the object we are testing [6, 14]. In fuzzy terms, this would mean to follow only the branch corresponding to the set where our value has a strictly positive membership degree. In the case of crisp decision trees, only one branch will have degree 1 and all the other branches will have degree 0, but for our fuzzy approach, we will consider all the branches for which the membership degree of our object to that particular set is nonzero. For each node, we will also have to keep in mind the membership of the object down the current path from the root of the node, where the membership is always 1, down to the current node. Again, in the case of crisp decision trees, the membership on the current path is always 1 (only one path and only one leaf have membership 1, all the other ones have

membership 0 and are therefore ignored). For simplicity, we will refer to the membership degree of an object to a certain set of attributes as the membership of the object to the corresponding tree node (or to the corresponding path in the tree). This membership can be computed gradually and needs to use an intersection operator for fuzzy memberships. In other words, considering that the attributes A_1, A_2, \dots, A_k were encountered from the root of the tree to a certain node (the node will in this case be on level k), below table shows the comparison between crisp and fuzzy decision trees [8, 9, 15] (Fig. 3).

The algorithm complexity of constructing a kind of minimum fuzzy decision tree is NP-hard. Construction of such trees needs considering a directed tree whose each edge links two nodes, the initial node, and the terminal node. The former is called the father node of the latter, while the latter is said to be the son node of the former. The node having not its father node is the root, whereas the nodes having not any son nodes are called leaves. The general matching strategy of fuzzy decision trees is described as follows:

1. Matching starts from the root and ends at a leaf along the branch of the maximum membership.
2. If the maximum membership at the node is not unique, matching proceeds along several branches.
3. The decision with maximum degree of truth is assigned to the matching result.

We illustrate this with an example. Consider a dataset [15] as in Fig. 4.

In the above figure, each column corresponds to a fuzzy subset defined as $X = \{1, 2, 3, \dots, 16\}$. In the last column, three symbols V , S , and W denote three sports to play on weekends: volleyball, swimming, and weightlifting, respectively. Now, keeping in mind the construction process as explained in previous pages, we define four attributes as follows:

Crisp Decision Trees	Fuzzy Decision Trees
Nodes are crisp subsets of X	Nodes are fuzzy subsets of X
If N is not a leaf and $\{N_i\}$ is the family off all sonnodes of N , then $\bigcup_i N_i = N$	If N is not a leaf and $\{N_i\}$ is the family off all sonnodes of N , then $\bigcup_i N_i \subset N$
A path from the root to a leaf corresponds to a production rule.	A path from the root to a leaf corresponds to a fuzzy rule with some degree of truth.
An example remaining to be classified matches only one path in the tree.	An example remaining to be classified can match several paths in the tree.
The intersection of subnodes located on the same layer is empty.	The intersection of subnodes located on the same layer can be nonempty.

Fig. 3 Comparison of crisp and fuzzy decision trees

No	Outlook			Temperature			Humidity		Humid Normal		Class		
	Sunny	Cloudy	Rain	Hot	Mild	Cool	Humid	Normal	Windy	Not windy	V	S	W
1	0.9	0.1	0.0	1.0	0.0	0.0	0.8	0.2	0.4	0.6	0.0	0.8	0.2
2	0.8	0.2	0.0	0.6	0.4	0.0	0.0	1.0	0.0	1.0	1.0	0.7	0.0
3	0.0	0.7	0.3	0.8	0.2	0.0	0.1	0.9	0.2	0.8	0.3	0.6	0.1
4	0.2	0.7	0.1	0.3	0.7	0.0	0.2	0.8	0.3	0.7	0.9	0.1	0.0
5	0.0	0.1	0.9	0.7	0.3	0.0	0.5	0.5	0.5	0.5	0.0	0.0	1.0
6	0.0	0.7	0.3	0.0	0.3	0.7	0.7	0.3	0.4	0.6	0.2	0.0	0.8
7	0.0	0.3	0.7	0.0	0.0	1.0	0.0	1.0	0.1	0.9	0.0	0.0	1.0
8	0.0	1.0	0.0	0.0	0.2	0.8	0.2	0.8	0.0	1.0	0.7	0.0	0.3
9	1.0	0.0	0.0	1.0	0.0	0.0	0.6	0.4	0.7	0.3	0.2	0.8	0.0
10	0.9	0.1	0.0	0.0	0.3	0.7	0.0	1.0	0.9	0.1	0.0	0.3	0.7
11	0.7	0.3	0.0	1.0	0.0	0.0	1.0	0.0	0.2	0.8	0.4	0.7	0.0
12	0.2	0.6	0.2	0.0	1.0	0.0	0.3	0.7	0.3	0.7	0.7	0.2	0.1
13	0.9	0.1	0.0	0.2	0.8	0.0	0.1	0.9	1.0	0.0	0.0	0.0	1.0
14	0.0	0.9	0.1	0.0	0.9	0.1	0.1	0.9	0.7	0.3	0.0	0.0	1.0
15	0.0	0.0	1.0	0.0	0.0	1.0	1.0	0.0	0.8	0.2	0.0	0.0	1.0
16	1.0	0.0	0.0	0.5	0.5	0.0	0.0	1.0	0.0	1.0	0.8	0.6	0.0

Fig. 4 Illustrative dataset [15]

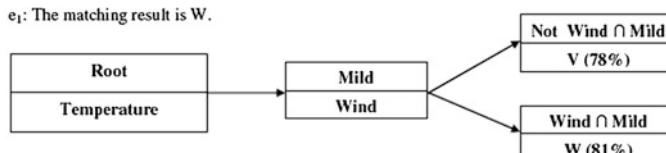
$$\text{Temperature} = \{\text{hot, mild, cool}\} \subset F(X)$$

$$\text{Outlook} = \{\text{sunny, cloudy, rain}\} \subset F(X)$$

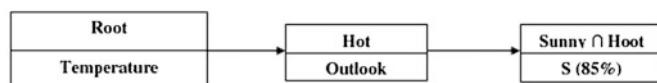
$$\text{Humidity} = \{\text{humid, normal}\} \subset F(X)$$

$$\text{Wind} = \{\text{windy, not_windy}\} \subset F(X)$$

A fuzzy decision tree for the root *Temperature* is given in Fig. 5. The percentage attached to each decision is the degree of truth of the decision. Consider two examples remaining to be classified



e₂: The matching result is S.

**Fig. 5** The process of matching in the fuzzy decision tree for temperature example

$$e_1 = (0.0, 0.6, 0.4; 0.3, 0.7, 0.0; 0.5, 0.5; 0.5, 0.5) \text{ and}$$

$$e_2 = (0.9, 0.1, 0.0; 0.8, 0.2, 0.0; 0.5, 0.5; 0.8, 0.2).$$

The process of matching in the fuzzy decision tree is shown in Fig. 5.

Keeping all the above discussions in mind, we can refine the traditional tree searchers. Prof. Ashit Kumar Dutta [13] has already highlighted these considerations. He has considered a simple search tree with one root node/start node and some descendent nodes. Consider a simple search path from root node/start node to some node in the tree. Accordingly, each node is labeled with only the terminal node of the path it denotes. Each child denotes a path that is a one-city extension of the path denoted by its parent. The most important searching method [16–18] that have analyzed so far includes DFS, BFS, Hill Climbing, A*, and IDA*. The depth-first search is a good idea if depth is less and breadth is more. In contrast, depth-first search is a bad idea if there are long paths, even infinitely long paths, that neither reach dead ends nor become complete paths. In those situations, we need alternative search methods. Breadth-first search works even in trees that are infinitely deep or effectively infinitely deep. On the other hand, breadth-first search is wasteful when all paths lead to the goal node at more or less the same depth. The breadth-first search is a bad idea if the branching factor is large or infinite, because of exponential explosion. Breadth-first search is a good idea when we are confident that the branching factor is small. We may also choose breadth-first search, instead of depth-first search, if we are worried that there may be long paths, even infinitely long paths, that neither reach dead ends nor become complete paths [13, 18]. We may be so uninformed about the search problem that we cannot rule out either a large branching factor or long useless paths. In such situations, we may want to seek a middle ground between depth-first search and breadth-first search. Search efficiency may improve spectacularly if there is a way to order the choices so that the most promising ones are explored at the earliest. In many situations, we can make measurements to determine a reasonable ordering. We have also learned about such methods that take advantage of such measurements; they are called heuristically informed methods such as Hill Climbing, A*, and IDA*. For better search efficiency, we can use the concept of shortest path. By the shortest path, we mean the path of which total distance is minimum [17, 18]. An obvious procedure is that one can find out all possible paths (if possible) and select the best one from them. For this, one can find out all possible paths by DFS or BFS and the search has to be continued until all paths are found. But in the maximum case, the size of the search tree is too large and consequently finding all possible paths is extremely rigorous. In many cases, we can improve the existing searching technique by using guesses about distances remaining, as well as facts about distances already accumulated. After all, if guess about distance remaining is good, then that guessed distance added to the definitely known distance already traversed should be a good estimate of total path length. Surely, it makes sense to work hardest on developing the path with the shortest estimated path length until the estimate is revised upward enough to make some other path be the one with the shortest estimated path length.

[6, 13, 15]. After all, if the guesses were perfect, this approach would keep us on the optimal path at all times. In general, however, guesses are not perfect, and a bad overestimate somewhere along the true optimal path may cause us to wander away from that optimal path permanently. Here, we make an interesting note that the underestimates cannot cause the right path to be overlooked. An underestimate of the distance remaining yields and underestimate of total path length u . Now, if we find a total path by extending the path with the smallest underestimate repeatedly, we need to do no further work once all partial path distance estimates are longer than the best complete far encountered [13]. We can stop because the real distance along a complete path cannot be less than an underestimate of that distance if all estimates of remaining distance can be guaranteed to be underestimates. While doing so, we need to remember that the closer an underestimate is to the true distance, the more efficiently we search. It is because, if there is no difference at all, there is no chance of developing any false movement. At the other extreme, an underestimate may be so poor as to be hardly better than a guess of zero, which certainly must always be the ultimate underestimate of remaining distance [6, 13, 15]. In fact, ignoring estimates of remaining distance altogether can be viewed as the special case in which the underestimate used is uniformly zero.

4 Conclusion

This paper highlights a method of interpreting decision tree searches by using fuzzy logic. This interpretation is labeled as characterization of fuzzy tree searches. Although a lot of issues are still up for discussion such as choosing the appropriate fuzzy operators and choosing the crisp partition on each attribute. By using this fuzzy characterization of searches, we get a new interpretation of the nodes, which help us to achieve greater accuracy when computing in uncertainties.

Acknowledgments Prof. Ashit Kumar Dutta, Prof. Richard Korf.

References

1. Mendonça, L.F., Susana M.V., Sousa, J.M.C.: Decision tree search methods in fuzzy modeling and classification. *Int. J. Approximate Reasoning* **44**(2), 106–123 (2007)
2. Pol, C.: Fuzzy decision trees. Lecture Notes. Available at www.cs.ubbcluj.ro/~gabis/DocDiplome/DT/FuzzyDT/fuzzy
3. Zadeh, L.A.: A theory of approximate reasoning. In: Hayes, J., Michie, D., Mikulich, L.I. (eds.) *Machine Intelligence*, vol. 9, pp. 149–194. Halstead Press, New York (1979)
4. Quinlan, J.R.: Induction on decision trees. *Mach. Learn.* **1**:81–106 (1986)
5. Quinlan, J.R.: Decision trees as probabilistic classifiers. In: *Proceedings of the Fourth, International Workshop on Machine Learning*, pp. 31–37 (1987)
6. Kahraman, C.: *Investment Analyses Using Fuzzy Decision Trees*. Lecture Notes. Istanbul Technical University, Department of Industrial Engineering, 34367, Macka, İstanbul, Turke

7. Zadeh, L.A.: Fuzzy sets. *Inf. Control* **8**, 338–353 (1965)
8. Zadeh, L.A.: Fuzzy logic and approximate reasoning. *Synthese* **30**, 407–428 (1975)
9. Cintra, M.E., Monard, M.C., Camargo, H.A.: Fuzzy DT—a fuzzy decision tree algorithm based on C4.5. In: National or Ibero-american conference publications-CBSF'2012: Congreso Brasileiro de Sistemas Fuzzzy, Natal, RN, Brasil (2012)
10. Jager, R.: Fuzzy logic in control. Ph.D. dissertation, Technische Universiteit, Delft (1995)
11. Janikow, C.Z.: Fuzzy decision trees: issues and methods. Department of Mathematics and Computer Science, University of Missouri-St. Louis
12. Jang, J.: Structure determination in fuzzy modeling: a fuzzy CART approach. In: Proceedings of the IEEE Conference on Fuzzy Systems, pp. 480–485 (1994)
13. Dutta, A.K.: Intuitionistic fuzzy search using dynamic programming. *J. Comput. Sci. Inform. Technol.* **3**(1), 1–6 (2006). ISSN 0973-4872 (Institute of Technology and Management, Gurgaon, Voyager)
14. Quinlan, J.R.: The effect of noise on concept learning. In: Michalski, R., Carbonell, J., Mitchell, T. (eds) *Machine Learning II*. Morgan Kaufmann, Burlington (1986)
15. Yuan, Y., Shaw, M.J.: Induction of fuzzy decision trees. *Fuzzy Sets Syst.* **69**, 125–139 (1995)
16. Bhalchandra, P.U., Khamitkar, S.D., Lokhande, S.N.: Pathology of trees, tree searches and their computational complexities. *JIKRCSA Int. J.* **2**(2) (2014)
17. Bhalchandra, P.U., Khamitkar, S.D., et al.: A review of contemporary research investigations in selected brute force and heuristics tree search algorithms. *IISHRE Int. J.* **2**(1) (2014)
18. Bhalchandra, P.U., et al.: Experimental analysis of selected constraint satisfaction problems using tree searching algorithms. In: Proceedings of ACM Sponsored ICCI, Mumbai (2014)

A Modified SOM-Based RBFN for Rotation Invariant Clear and Occluded Fingerprint Recognition

Sumana Kundu and Goutam Sarker

Abstract In this paper, a modified radial basis function network (RBFN) based on self-organization mapping (SOM) has been designed and developed for rotation invariant clear as well as occluded fingerprint recognition. The SOM-based RBFN learns different fingerprint images and performs subsequent rotation invariant recognition of clear and occluded images. The system is efficient, effective, and fast. Also, the performance evaluation of the system is substantially moderate.

Keywords SOM · RBFN · BP learning · Occlusion · Holdout method · Accuracy · Precision · Recall · F-score

1 Introduction

Fingerprint recognition is one of the vital biometrics methods that has been widely used in various applications because of its reliability and accuracy in the process of recognizing and verifying a person identity.

Most of the fingerprint recognition methods are based on feature (minutiae) extraction and minutiae matching [1, 2]. In [3], a fingerprint identification technology was represented by minutiae feature extraction using back propagation (BP) algorithm.

A three-rate hybrid Kohonen neural network (TRHKNN) was proposed in [4] for distorted fingerprint image processing in conditions of wide variation in degree of distortion. This proposed TRHKNN is capable to identify the distorted image of fingerprint and restore the undistorted image of fingerprint.

S. Kundu (✉) · G. Sarker

Computer Science and Engineering Department, NIT, Durgapur, India
e-mail: sumana.kundu@yahoo.co.in

G. Sarker

e-mail: g_sarker@ieee.org

A fingerprint classification method was proposed based on self-organizing maps (SOM) in [5]. The concept of “certainty” was introduced and used in the modified algorithms, and the size of the network was increased.

In our previous works [6, 7], we have developed a RBFN with Optimal Clustering Algorithm (OCA), and in [8], we have developed a RBFN with Heuristic-Based Clustering (HBC) for clear and occluded location invariant fingerprint recognition. However, these systems are not adapted to recognize rotation invariant fingerprints excepting 90° and 180° rotation.

In the present paper, a self-organizing mapping-based modified RBFN is used for clear and occluded rotation invariant fingerprint identification.

2 Theory of the Operation

In this approach, a feature mapping network that is self-organization network [5, 9] with Kohonen’s learning is used to form groups of the preprocessed input data set which are taken as input of the radial basis function network (RBFN). The SOM can construct the two-dimensional feature map (here, we use 15×15 feature map) from which the number of clusters can be figured out directly. It forms groups of different qualities of fingerprints of each person and angle (person-angle). The mean “ μ ” and standard deviation “ σ ” of each cluster formed by SOM with approximated normal distribution output function are used for each basis unit of RBFN. Then, BP Learning Algorithm is presented which classifies the “person fingerprint-angle” into “person fingerprint.” In the output layer, using BP learning, we get the optimal weights. Here, we set the number of inputs equals to the number of fingerprint images present in the training database, while the number of outputs sets to the number of classes and the number of hidden units is equal to the number of cluster formed by SOM based on qualities of fingerprints of each “person-angle.” Hence, the modified RBFN with BP network classifier is used for fingerprint identification.

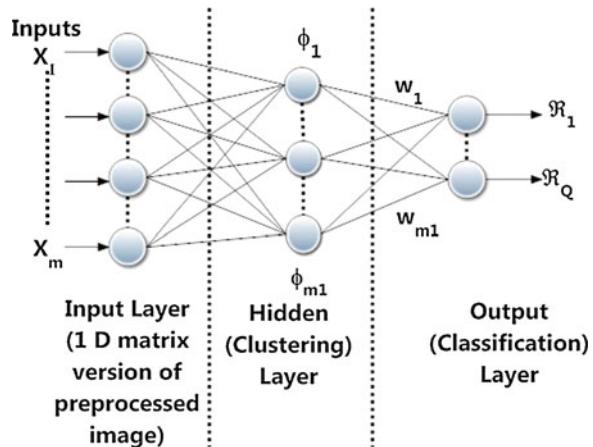
2.1 Preprocessing

The finger print images of training and test database have to be preprocessed before learning and recognition. The different steps in preprocessing are as follows:

1. Conversion of RGB fingerprints images to grayscale images.
2. Removal of noise from images.
3. Deblurring of the images.
4. Background elimination.
5. Conversion of grayscale images into binary images.
6. Image normalization.
7. Conversion of binary images into 1D matrix.

This set is the input to the self-organizing network.

Fig. 1 Modified RBFN architecture



2.2 Radial Basis Function Network (RBFN)

The RBFN of Fig. 1, [6–8, 10, 11], consists of three layers, namely an input layer for fingerprint pattern presentation, a hidden (clustering) layer containing “basis units,” and an output (classification) layer. The clustering outputs (mean μ , standard deviation σ , and corresponding approximated normal distribution output functions) are used in “basis units” of RBFN. Thus, SOM is the first phase learning, and using BP learning [6–8, 12, 13], we get the optimal weights in the BP network, which is the second phase of learning. The hidden layer uses neurons with RBF activation functions describing local receptors. Then, output node is used to combine linearly the outputs of the hidden neurons. Here, “ m ” denotes the number of inputs while “ Q ” denotes the number of outputs.

2.3 Identification Learning

The training database contains finger prints of 4 different people. For each person’s finger print, 5 different qualities of finger prints and also 3 different angular (0° , 90° , and 180°) finger prints are there. After preprocessing, all the finger print images are fed as input to SOM network. When the SOM-based RBFN network has learned all the different qualities of fingerprints of all the angles (0° , 90° , and 180°) for all the different persons, the network is ready for recognition of learned fingerprint images (refer to Figs. 2 and 3).

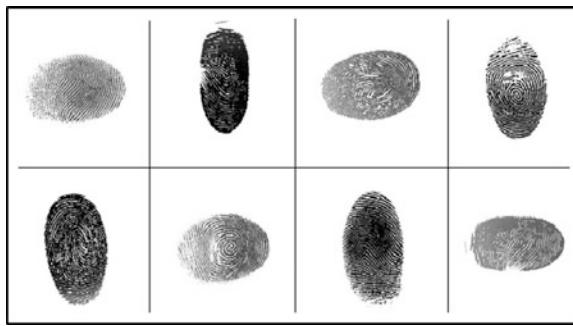


Fig. 2 Set of few training fingerprint images for recognition

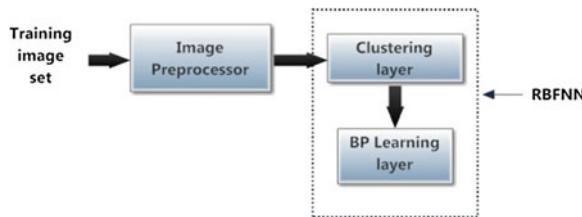


Fig. 3 Block diagram of learning identification

2.4 Identification with Test Image

The labeled test image set consists of 4 different person's (same as training data set) clear and occluded finger prints of various different angles, and the qualities are different with respect to training set. It also contains some unknown fingerprints of various qualities and angles. The test image from the test image set is fed as input to the preprocessor. The preprocessed pattern is fed as input to the previously trained BP network. If any of the 4 outputs of the network is active high, then the corresponding finger print is identified or recognized.

If the test fingerprints are different angular fingerprints other than (0° , 90° , and 180°), then we calculate the angle of rotation and rotate that corresponding fingerprint by this angle and then feed this rotated fingerprint to the previously trained BP network. In this case also, if any of the 4 outputs of the network is active high, then the corresponding finger print is identified or recognized (refer to Figs. 4 and 5).

Fig. 4 Set of few clear test images with or without rotation

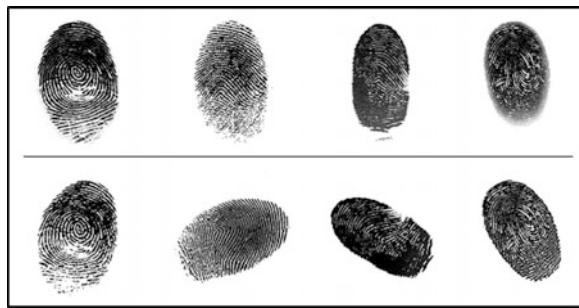
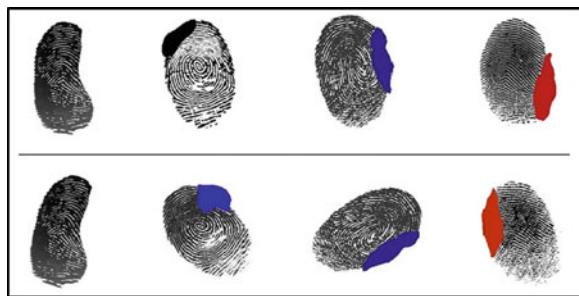


Fig. 5 Set of few occluded test images with or without rotation



3 Result and Performance Analysis

We have used the training and test data set for fingerprints samples from FVC database (<http://www.advancedsourcecode.com/fingerprintdatabase.asp>).

3.1 Classifier's Performance Evaluation

We use holdout method [14] to evaluate the performance of the classifier (refer Tables 1 and 2; Figs. 2, 4, and 5).

From Fig. 6 mentioned confusion matrix, if there are only two classes (say X and Y), then the accuracy, precision, recall, and F-score are defined as follows:

$$\text{Accuracy} = \frac{a + d}{a + b + c + d} * 100 \quad (1)$$

$$\text{Precision} = \frac{a}{a + b} \quad (2)$$

$$\text{Recall} = \frac{a}{a + c} \quad (3)$$

Table 1 Accuracy of the classifier (Holdout method) and fingerprint learning time (seconds)

Type of the fingerprint	Learning time (in seconds)			
	Accuracy (%)	Training time	Performance evaluation time ^a	Total (learning) time
Single clear fingerprints	96.29	90.17	0.17	90.34
Single occluded fingerprints	93.33	90.17	0.17	90.34
Single clear and rotated fingerprints	90.00	90.17	0.18	90.35
Single occluded and rotated fingerprints	85.00	90.17	0.17	90.34

^a We take 10 test samples from test data set of fingerprints to find out performance evaluation time

Table 2 Precision of the classifier (Holdout method)

Type of the fingerprint	Person 1	Person 2	Person 3	Person 4	Unknown person
Clear fingerprints	1	1	0.90	1	0.92
Occluded fingerprints	1	1	0.93	0.84	0.91
Clear rotated fingerprints	1	1	1	0.78	0.75
Occluded rotated fingerprints	1	1	0.78	1	0.63

		Actual Class		
		X	Y	
Predicted Class	X	a	b	
	Y	c	d	

Fig. 6 Confusion matrix (2 class)

$$F\text{-score} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (4)$$

3.1.1 Experimental Results

The proposed system was made to learn on a computer with Intel Core 2 Duo E8400, 3.00 GHz processor with 4 GB RAM and Windows 7 32-bit Operating System.

From Table 1, we find the *accuracy* of the classifier with holdout method. From Tables 2, 3, and 4, *precision*, *recall*, and *F-score* metrics for different clear, occluded, rotated fingerprints clarify the performance of each class. Again in Table 1, the proposed approach shows overall low performance evaluation time (<0.02 s for each fingerprint) for the standard test data set. Hence, the proposed approach shows improvement in terms of accuracy and learning as well as

Table 3 Recall of the classifier (Holdout method)

Type of the fingerprint	Person 1	Person 2	Person 3	Person 4	Unknown person
Clear fingerprints	0.93	1	1	1	0.89
Occluded fingerprints	0.96	0.96	1	1	0.74
Clear rotated fingerprints	0.88	1	1	0.88	0.75
Occluded rotated fingerprints	0.63	1	0.88	0.88	0.88

Table 4 F-score of the classifier (Holdout method)

Type of the fingerprint	Person 1	Person 2	Person 3	Person 4	Unknown person
Clear fingerprints	0.96	1	0.95	1	0.90
Occluded fingerprints	0.98	0.98	0.96	0.91	0.82
Clear rotated fingerprints	0.94	1	1	0.83	0.75
Occluded rotated fingerprints	0.77	1	0.83	0.94	0.73

performance evaluation time which is comparably appreciable than the present existing systems [1–5] mentioned in the Sect. 1. So, the main contribution of this paper is to perform accurate fingerprint identification.

4 Conclusion

A self-organizing mapping (SOM)-based modified RBFN has been designed and developed for fingerprint recognition. This system is able to identify rotation invariant clear as well as occluded finger prints. The performance measurement in terms of accuracy, precision, recall, and F-score with holdout method is moderately high for fingerprint images. Also, the learning as well as performance evaluation time is moderately low for different types of fingerprints. Due to the application of our new SOM-based RBFN, the present system is efficient, effective, and faster compared to any other conventional fingerprint identification technique.

References

1. Bana, S., Kaur, D.: Fingerprint recognition using image segmentation. *IJAEST* **5**(1), 012–023 (2011)
2. Kovács-Vajna, Z.M.: A fingerprint verification system based on triangular matching and dynamic time warping. *IEEE Trans. PAMI* **22**(11), 1266–1276 (2000)
3. Chatterjee, A., Mandal, S., Rahaman, G.M.A., Arif, A.S.M.: Fingerprint identification and verification system by minutiae extraction using artificial neural network. *JCIT* **1**(1), 12–16 (2010)
4. Astrov, I., Tatarly, S., Tatarly, S.: Fingerprint recognition for varied degrees of image distortion using three-rate hybrid kohonen neural network. In: *ICALIP*, pp. 363–369 (2008)

5. Turky, A.M., Ahmad, M.S.: The use of SOM for fingerprint classification. In: International Conference on Information Retrieval and Knowledge Management, pp. 287–290 (2010)
6. Kundu, S., Sarker, G.: A modified radial basis function network for occluded fingerprint identification and localization. IJCITAE 7(2), 103–109 (2013)
7. Sarker, G., Kundu, S.: A modified radial basis function network for fingerprint identification and localization. In: International Conference on Advanced Engineering and Technology, pp. 26–31 (2013)
8. Kundu, S., Sarker, G.: A modified RBFN based on heuristic based clustering for location invariant fingerprint recognition and localization with and without occlusion. In: IEEE's International Conferences for Convergence of Technology (2014)
9. Yegnanarayana, B.: Artificial Neural Networks, pp. 223–228. PHI Learning Private Limited, Delhi, (1999)
10. Aziz, K.A.A., Ramlee, R.A., Abdullah, S.S., Jahari, A.N.: Face detection using radial basis function neural networks with variance spread value. In: International Conference of Soft Computing and Pattern Recognition, pp. 399–403 (2009)
11. Ashok, J., Rajan, E.G.: Off-line hand written character recognition using radial basis function. Int. J. Adv. Network. Appl. 2(2), 792–795 (2011)
12. Liu, X., Geng, G., Wang, X.: Automatically face detection based on bp neural network and bayesian decision. In: 2010 Sixth International Conference on Natural Computation, pp. 1590–1594 (2010)
13. Sarker, G.: A back propagation network for face identification and localization. In: International Conference on Recent Trends in Information Systems, pp. 24–29 (2011)
14. Pudi, V.: Data mining. Oxford University Press India, Oxford (2009)

An Unsupervised OCA-based RBFN for Clear and Occluded Face Identification

Dhananjoy Bhakta and Goutam Sarker

Abstract We present an automatic face identification system using an unsupervised optimal clustering algorithm (OCA)-based RBF network. In this present system, we propose a completely unsupervised clustering algorithm for training of the RBF network in which the system automatically searches for suitable threshold to perform natural clustering. This system performs the identity of a person irrespective of different facial expressions, poses, and partial occlusions. Experimental results show that the performance of the system in terms of accuracy, precision, recall, and F-score is moderately high. At the same time, the total learning time as well as performance evaluation time is moderately low.

Keywords UOCA · RBFN · Occlusion · Accuracy · Precision · Recall · F-score

1 Introduction

Face recognition is an interesting and successful application of image processing, pattern recognition and computer vision which classifies a known or an unknown person and within the known category validates the identity of the person. Many methods have been proposed [1, 6, 9] for face identification tasks among which neural network approaches achieve higher accuracy. Due to very fast learning speed, many researchers used radial basis function (RBF) network for this purpose [7]. In most of the real-time face recognition system, it is necessary to overcome some difficulties such as identifying a face with different poses, highly variable expressions, occlusions where occlusion poses a significant obstacle due to the

D. Bhakta (✉) · G. Sarker

Computer Science and Engineering Department, NIT Durgapur, Durgapur, India
e-mail: bhaktadhananjoy@gmail.com

G. Sarker

e-mail: g_sarker@ieee.org

unpredictable nature of the error incurred by occlusion, which may affect any part of the image and may be arbitrarily large in magnitude [8].

In our previous work [3, 4], we have developed RBFN with optimal clustering and BP Network. These systems were unable to deal with occluded faces. Also, the localizations were not location invariant. In one of our previous works [5], a multilayer BP network was used for face identification and localization. The system was unable to handle occlusion.

Gao [10] presents a novel string-based face recognition approach to deal with partial occlusion problem in which face image treated as a string of local structural information with spatial information of a face image by grouping the relational organization of intermediate-level features (line segments). The recognition rate of the system is varying from 76 to 96 % in occluded face and 25–96 % in different lighting conditions and facial expression changes.

In [11], the author proposed a new fractal-based face recognition system which mainly deals with expressions and occlusions. The system shows that the recognition rate in highly screamed expression is 60–84 % but in normal case, it is 96–99 % while in case of occlusion, it is about 85–97 %.

From the above discussion, it is clear that developing a face identification technique which works efficiently with both clear and occluded images is yet to be an active research area. This paper mainly considers identifying a face with different facial expressions, different poses, and partial occlusions. We implement an efficient face identification system with an unsupervised OCA (UOCA) [2]-based modified RBFN. In order to obtain natural grouping of different expressions of each “person-angle,” UOCA is applied to find out the mean and standard derivation of each group of expressions for all “person-angle” which is used in the middle layer units of RBFN. Then, back propagation (BP) learning is presented to train the RBFN for performing identification.

To evaluate the performance of our identifier, we used holdout method from several other approaches [12] where the training dataset (labeled instances) is used to build the identifier and test dataset is used to test the performance of the constructed identifier.

In Sect. 2, UOCA, BP Learning and RBFN, and algorithm used in the system are briefly introduced. Overview of the present system is discussed in Sect. 3. Experimental results based on database from FEI (fei.edu.br/~cet/facedatabase.html) are reported in Sect. 4. Finally, Sect. 5 presents the conclusion.

2 Theory of Operation

2.1 Unsupervised Optimal Clustering Algorithm (UOCA)

To perform natural clustering on face images, we used UOCA [2] in which the system automatically searches for the suitable conceptual affinity or threshold (measurement of the degree of similarity among the members of the same cluster

and that of dissimilarity among members of different clusters). Unlike, k-means or supervised OCA [2], this clustering technique is free from choosing any arbitrary value of k or threshold which might produce some undesirable cluster. In ordinary OCA, the threshold has to be supplied by the user to govern the desire number of cluster. But the unsupervised OCA finds threshold automatically by the system. This new clustering technique is used for grouping as “person-angle” pattern. For example, person 1 with different expressions is grouped as the same person with different face orientations (front view, left/right side view). This clustering technique invokes the ordinary OCA [2–4] algorithm repetitively. The UOCA can be stated as follows:

Input: A database containing n pattern points

Output: A set of natural number of clusters with k different clusters element $C_1, C_2, C_3, \dots, C_k$

Steps:

1. *begin*
2. *find mean of “n” pattern points. Also find the distance of the furthest pattern point from the mean and assign it equal to “r”*
3. *for $Th = d$ to r (d = length of a pattern point) perform OCA Clustering*
4. *compute the number of clusters $N(Th)$*
5. $Th = Th + d$
6. *Plot $N(Th)$ vs. Th*
7. *find the region(s) where $N(Th)$ is almost independent of variations of Th . If the no. of cluster formation maximally varies by 2 for threshold variation of Th_1 to Th_2 then take the average of Th_1 and Th_2 (= Th^l). Let almost equal number of clusters be $N^l(Th^l)$ for $Th = Th^l$*
8. *if there are a number of such regions, compute all such Th^l and find the Th^l which is closest to the origin, i.e., which produces maximum number of clusters*
9. *exhibit that $N^l(Th^l)$ number of clusters formed with threshold value Th^l*
10. *end for*
11. *end*

The time complexity of the algorithm is $O(k*(r/d)*n)$, where “ k ” is the number of clusters formed, “ r ” is radius of pattern set, “ d ” is the length of a pattern point, and “ n ” is the no. of pattern points in the dataset.

2.2 Back Propagation(BP) Learning

One of the common methods for training artificial neural network is BP Learning [3–5]. In this learning, first we set the target output. Now, the output of the network can be used to update the weight of the different links based on the error, i.e., difference between the network output and target output. So, the network output is back propagated. This identifies a person with different angular views correctly. For ex., person 1 with different angular views is recognized as person 1.

2.3 RBF Neural Network

The RBF neural network is a feed-forward neural network, which accomplishes an input–output nonlinear mapping by a linear combination of nonlinearly transformed input pattern. The result of clustering, that is a set of centers and standard deviation are used to train the middle layer (hidden) units of RBFN and BP learning is used to adjust synaptic weight between hidden layer units and output layer units. The learning saturates, when no further presentation of image in overall input can change the weight of any link, i.e., $(w_{ij})_{t+1} \approx (w_{ij})_t$. Change in weight $\Delta w_{ij} = w_{ij} + \eta \cdot \text{error}$, where $(w_{ij})_t$ is weight vector between j th to i th layer at t th iteration and “ η ” is small fraction called the “rate of learning.” We have assumed $\eta = 0.1$.

3 Overview of the System and Approach

The proposed system is described as follows:

3.1 Preprocessing

The first stage of our face identification system is preprocessing where the face images have to be converted into gray-scale face images. Then, we used smoothing filter to reduce noise which may produce blurring effect in images. We apply median filter to get back the de-blurred images. Now, we remove background images and convert it into binary images. The images then compressed by replacing block of pixels with the mode value of pixel intensities, which normalizes the all images of 640×480 into 128×96 . Finally, the compressed images are converted into 1D column matrix which is fed as input to the system.

3.2 Modified RBF Network

In this stage, at first, UOCA performs natural clustering on the preprocessed face images which performs grouping of “person-angle” irrespective of facial expressions. Now, the clustering result as a set of mean and standard deviation is fed as input to the hidden layer of BP network. Then, the system constantly adjusts the weight of the links each time a new image is presented in overall input until learning is saturates. This identifies a person with different angular views correctly. For example, person 1 with different angular views is recognized as person 1.

After training, the modified RBFN gives high-output value for known faces and low-output value for unknown faces. To distinguish known faces from the unknown faces, we need to set a threshold value [4]. The corresponding output value above threshold is considered as corresponding known person.

4 Results and Performance Analysis

The platform under which the system was made to learn was given as follows: Intel Core 2 Duo CPU E8400@3.00 GHz, 4 GB RAM, and 32-bit operating system.

Experimental results are carried out on the FEI database (fei.edu.br/~cet/face-database.html) [13]. The first phase of learning UOCA takes around 28 min, and the second phase of learning BP learning takes only 9.41 s for the training dataset entitled “TrainingFace4”.

For our system, we select five different faces each of which have three different views each with three different expressions (normal, smile, and angry) as training example set. Similarly, we create test dataset from the training example set and other dataset by changing their expressions and/or adding some occlusion. Performance evaluation was done with 90 clear as well as occluded facial instances.

From the below-mentioned confusion matrix (Fig. 1), accuracy, precision, recall, and F-score of the classifier can be defined as follows:

$$\text{Accuracy} = \frac{a + d}{(a + b + c + d)} \quad (1)$$

$$\text{Precision} = \frac{a}{(a + b)} \quad (2)$$

$$\text{Recall} = \frac{a}{(a + c)} \quad (3)$$

$$F\text{-score} = \frac{2 \times \text{recall} \times \text{precision}}{(\text{recall} + \text{precision})} \quad (4)$$

The experimental results are shown in Tables 1, 2, and 3 in terms of accuracy, learning time, precision, recall, and F-Score. In Tables 2 and 3, P1, P2, P3, P4, and P5 indicate person 1, person 2, person 3, person 4, and person 5, respectively. From the result, we can see that the system gives 98.75 % accuracy for clear faces where it is 94.17 % for occluded faces (up to 50 % occlusion in face images) and in case of learning, total time taken by 90 faces are less than 30 min. Also, the precision, recall, and F-score are lies between 96 and 100 % for each person (with clear faces) and 89–100 % for each person (with occluded faces) which performs better than ordinary OCA-based classifier [14].

		Actual Class	
		C	C̄
Predicted class	C	a	b
	C̄	c	d

Fig. 1 Confusion matrix (2 Class)

Table 1 Accuracy and learning time of the classifier

Type of the face	Accuracy (in %)	Learning time (in sec.)		
		Training time	Test time	Total (learning) time
Clear faces	98.75	1740.41	2.889	1743.29
Occluded faces	94.17	1740.41	2.394	1742.80

Table 2 Precision, recall, and F-score of the classifier for clear faces

Performance metric	<i>P</i> 1	<i>P</i> 2	<i>P</i> 3	<i>P</i> 4	<i>P</i> 5
Precision	0.98	1	1	0.96	1
Recall	1	0.98	1	1	0.96
F-score	0.99	0.99	1	0.98	0.98

Table 3 Precision, recall and F-score of the classifier for occluded faces

Performance metric	<i>P</i> 1	<i>P</i> 2	<i>P</i> 3	<i>P</i> 4	<i>P</i> 5
Precision	0.96	1	1	0.89	0.92
Recall	0.96	0.92	0.95	0.96	0.95
F-score	0.96	0.96	0.97	0.92	0.94

5 Conclusion

An efficient face identification system has been designed and developed to identify faces irrespective of facial expression, pose, and partial occlusion using an unsupervised OCA-based modified RBFN. The system successfully identifies faces with highly variable expression and up to 50 % occlusion. Due to the application of completely unsupervised clustering techniques, the system automatically finds the suitable threshold to form natural number of cluster. Therefore, with the help of modified RBFN and unsupervised OCA technique, the accuracy, precision, recall, and F-score of the system are moderately high for both clear as well as occluded face images. Also, the total learning time is moderately low.

References

1. Yang, M.H., Kriegman, D.J., Ahuja, N.: Detecting faces in images: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(1), 34–58 (2002)
2. Sarker, G.: An unsupervised natural clustering with optimal conceptual affinity. *J. Intell. Syst.* **19**(3), 289–300 (2010)
3. Sarker, G., Roy, K.: A modified RBF network with optimal clustering for face identification and localization. *Int. J. Adv. Comput. Eng. Networking* **1**(3), 30–35 (2013)
4. Bhakta, D., Sarker, G.: A radial basis function network for face identification and subsequent localization. In: International Conference on Computer Science and Information Technology (ICCSIT), pp. 1–6 (2013)

5. Sarker, G.: An optimal back propagation network for face identification and localization. *Int. J. Comput. Appl. (IJCA)*, ACTA Press, **35**(2) (2013)
6. Wu, H., Chen, Q., Yachida, M.: Face detection from color images using fuzzy pattern matching method. *IEEE Trans. Pattern Anal. Mach. Intell.* **21**(6) (1999)
7. Moody, J., Draken, C.J.: Fast learning in network of locally tuned processing units. *Neural Comput.* **1**, 281–294 (1989)
8. Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Ma, Y.: Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(2), 1–18 (2009)
9. Ali, S.: Novel fast and efficient face recognition technique. *Int. J. IT Eng. Appl. Sci. Res. (IJIEASR)* **1**(1), 5–9 (2012)
10. Gao, W.C.Y.: Recognizing partially occluded faces from a single sample per class using string-based matching LNCS 6313, pp. 496–509. Springer, Berlin (2010)
11. Marsico, M.D., Nappi, M., Riccio, D.: FARO: Face recognition against occlusions and expression variations. *IEEE Trans. Syst. Man Cybern.—Part A: Syst. Humans* **40**(1), 121–132 (2010)
12. Pudi, V., Radhakrishna, P.: Data Mining. Oxford Press, India (2009)
13. Thomaz, C.E., Giraldi, G.A.: A new ranking method for principal components analysis and its application to face image analysis. *Image Vis. Comput.* **28**(6), 902–913 (2010)
14. Bhakta, D., Sarker, G.: A rotation and location invariant face identification and localization with or without occlusion using modified RBFN. In: IEEE 2nd International Conference on Image Information Processing (ICIIP-2013), pp. 533–538

Analysing the Impact of Human Behaviour over Machine Learning

Shailendra Singh Yadav and Bhupendra Verma

Abstract Influenced by the human learning, neural network promises to provide computer science discipline a breakthrough that rivals anything it has yet witnessed. Once neural networks are trained properly, they can replace many human functions in targeted areas. This study deals with an important step in that journey by analysing the impact of training data set drawn from the actual learning process of a human being, on artificial neural network learning. Use of kid's handwriting samples as training samples has been used to demonstrate its effect on neural network performance as it has maximum variance that helps neural network to exhaustively explore input space. The effect of training samples obtained from kid's handwriting has been analysed on handwritten character identification and compared with the training samples obtained from young subjects. From the study, it is concluded that neural networks trained by kid's writing data sample are able to recognize young subject writing test sample inputs with an increased average recognition rate of 13 % as compared to neural networks trained from young subject writing data. Similarly, neural networks trained by kid's writing are able to recognize kid's writing test sample inputs with an increased average recognition rate of 38.7 % as compared to neural networks trained from young subject samples.

Keywords Human behaviour • Offline character recognition • Profile information • Feed-forward BPN • Kid's data set

S.S. Yadav (✉)

Computer Science and Engineering Department, S.R.G.I, Jhansi, India
e-mail: shailendra.s.yadav@gmail.com

B. Verma

Director TIT Excellence, Bhopal, India
e-mail: bkverma3@gmail.com

1 Introduction

Properties of data samples may directly influence the learning of artificial neural network as network will be exposed to problem space only through the data samples [1]. This paper presents the study of effect of data on the learning of neural networks. Case study is presented in character recognition area to improve its performance with the help of training data samples collected from kid's handwriting to incorporate maximum diversity and to train network from the data generated from the responses of human brain during its actual learning phase. Studies have shown that brain responds to new things very rapidly, i.e. the novelty-seeking property of the human brain [2–4, 29].

In this paper, an attempt has been made to resolve the issues and problems of the offline character recognition system. The general handwritten character recognition system consists of the following three phases, i.e. pre-processing, feature extraction and classification. In pre-processing stage, input character image is learned, so that it can be used directly and efficiently by the feature extraction stage and finally by classification stage to classify the input into specified class based on the features. In this research work, profile information is used as feature extraction tool.

In this work, emphasis has been given to train the network by the kid's handwriting data set. The kids here are small children whose age lies between 4 and 6 years. The idea of considering samples collected from small children is human behaviour, i.e. response of human mind at this learning stage could be used to enhance the learning of network. When a small child tries to draw the alphabets as similar to the characters, at that instant his writing has maximum variance as it actually shows the output of human neural network. Since the artificial neural network simulates the human behaviour and the underlying principle of learning of artificial neural network is same as it actually takes place in humans, therefore, kid's data set is chosen to provide sufficient input to the network so that it can completely explore all the input data set space and the required generalization and fault tolerance can be achieved by the network after training. The usefulness of kid's data set for the training of the network is shown by the excellent results achieved during the testing phase.

The remainder of this paper is organized as follows: Sect. 2 briefly overviews the character recognition system. The pre-processing stage is briefly presented in Sect. 3. Section 4 describes the feature extraction phase, Sect. 5 reports the working of classifier, Sect. 6 describes the results, and Sect. 7 concludes the work.

2 Character Recognition System

The character recognition systems can be used significantly in the advancement of the automation process and can improve the way in which human and machine interact in various applications such as office automation, cheque verification in

banking sector, survey items collection, and mail sorting in post office, business and data entry applications [5–7]. The different approaches covered under the general term character recognition fall into either the online or the offline category, each having its own hardware and recognition algorithms. Offline recognition is performed, after the writing or printing is completed.

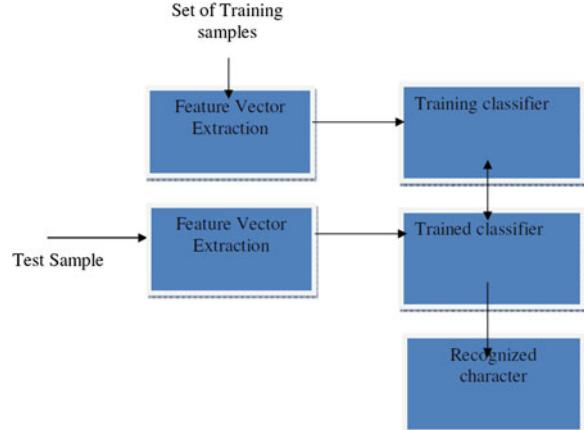
Before proceeding with the experiments and its details, let us have an idea of how any recognition system works. The main objectives of this system are to separate out the individual character and employ character classification on it. It is worth noting that performance of character classification can be boosted up by using ideal individual character segmentation. Character recognition system may use local and/or global descriptors for features representing characters. There are always advantages and disadvantages with one type of descriptor over another. In local feature extraction, features need to be extracted from the points that give maximum entropy on behalf of character. The feature locations are features extracted from feature points. The model in Fig. 1 can be applicable to any recognition system. The model consists of general steps, which is a need of any recognition system. The model will be suggested by me after reading various research papers related to the recognition system. The concluding steps in this model can be used by most of the systems [8–11].

Offline character recognition system consists of the following stages:

- (i) Pre-processing
- (ii) Feature extraction
- (iii) Classification

The steps of pre-processing and feature extraction are repeated for both training data set and testing data set. The model suggested above is a generalized model for any recognition system.

Fig. 1 Generalized character recognition system



3 Pre-processing Phase

Every human has different styles of writing, so there is a collection of various diversified characters, which is not a single-step process to recognize a character. In pre-processing step, the redundancy in data is removed such as unwanted edges and unwanted marks, which we call it as a noise. The raw data are subjected to a number of preliminary processing steps or stages to make it usable in the descriptive stages of character analysis process. Pre-processing [12, 13] aims to produce data that is easy for any OCR systems to operate accurately and precisely. The pre-processing steps include binarization and removal of any unwanted noise in an image.

Binarization involves in the process of converting grey images into black-and-white images (binary) which depend on the threshold value of any image. The input image can be of class uint8, uint16, single, int16 or double and must not be sparse. The process is used to convert pixel value into inverted form of an image to perform various operations over it.

The process reduces the noise from the image, which means that a small portion of unwanted image, which is not required, should be avoided with the help of finding mean area value of each character and morphing the noise values from the processed images that have lesser area than the threshold values. As in this work, the kid's data set that is used to train neural network contains 260 samples. But before this step, it was showing 292 characters (containing some noise in the data set). This is a very important process for testing and training data set to get the actual values of data set.

4 Feature Extraction Phase

Feature extraction is a method, which evaluates the unique identity of any object or pattern. In feature extraction stage, each character is represented as a feature vector, which becomes its identity. The major goal of feature extraction [14, 15] is to extract a set of features, which maximizes the recognition rate with the least amount of elements.

Feature extraction is needed to reduce the amount of data into feature vector. For example, any medical X-ray image always contains millions of pixels and that have to be extracted as a feature. The aim of feature extraction process is to distil or filter the information into a more concentrated and manageable form which is also known as feature vector of an image. The tool that is used to extract maximum feature of an alphabet is a need of any classifier for its effective training. In case of handwritten character, it is difficult to obtain these features due to divertive or cursive nature of handwriting as it contains high degree of variability and imprecision.

Feature extraction methods are based on three types of features:

1. Statistical
2. Structural
3. Global transformations and moments.

This research work used statistical feature extraction tool to extract features, from testing and training data set.

4.1 Statistical Feature

Representation of an isolated character image by statistical distribution points should take care of style variations to some extent. The profile information is used as statistical feature extraction tool. The profile [16] counts the number of pixels (distance) between the boundary position pixel of the character image of 15×15 pixel size and the edge of the character. The profiles describe well the external shapes of characters and allow distinguishing between letters, such as “p” and “q”. Since the profiles depend on the image dimension, the features are made scale independent by normalizing the profiles to equal bins at each axis to have an equal number of elements for all characters. The profiles are taken at 4 positions:

1. Top
2. Bottom
3. Left hand
4. Right hand

It is illustrated in Fig. 2 [16] in the next section.

4.2 Representation of a Profile

Here, it is a diagrammatic representation of a single character, which shows how the particular character represents its profile information. It was well defined by the diagram that profile allows distinction between characters such as “p” and “q”. Figure 2 shows how we should apply profile information to all characters, which

Fig. 2 Profile of any characters



would basically calculate the distance of first pixel position from any boundary position pixels. The distance from all four positions, i.e., top, bottom, right, left, is shown in the diagram for an alphabet “a”.

The profile extracted from each individual is a collection of distance value of each character pixel from all four directions.

5 Classification and Recognition

Object recognition is based on assigning classes to objects, and the device that does this assignment is called classifier [17]. Classification is the process, which can be used for both testing and training purposes. It can also be defined as an artificial brain, which has been trained by different objects to recognize things. There can be many methods for building a classifier such as SVM, HMM [18–20] and neural network. Here, feed-forward neural network has been used as a classifier [21–23, 28].

5.1 Data set Representation

This section describes the data set used for training purpose of our neural network. Data set is a diversified collection of various styles of character writing. So, two data sets have been taken, i.e. one is kid’s handwriting having age between 4 and 6 years and other is adult’s having age more than 18. The key role of data set and its performance have been described in the result section. Two data sets have been taken for the experiments, i.e. normal writing and kid’s writing. The data set plays a vital role in any pattern recognition application. The kid’s data set gives an encouraging result and can be continued as a research area in future for pattern recognition application mainly for character recognition.

5.2 Feed-Forward Neural Network

The multilayer feed-forward neural network has been used with back-propagation learning algorithm [24, 25]. Back-propagation is a systematic method of training a network, which is used to extend gradient descent-based delta learning rule, commonly known as back-propagation rule [26].

The multilayer distinguishes itself from one or more hidden layers, whose computation nodes are correspondingly called neurons or units [27]. The function of hidden neurons is intervening between the external input and the network output in some useful manner. The input vectors are fed forward to 1st hidden layer and this pass to 2nd hidden layer and so on until the last layer, i.e., output layer, which gives actual network response. The architecture of the network having two layers is illustrated in Fig. 3.

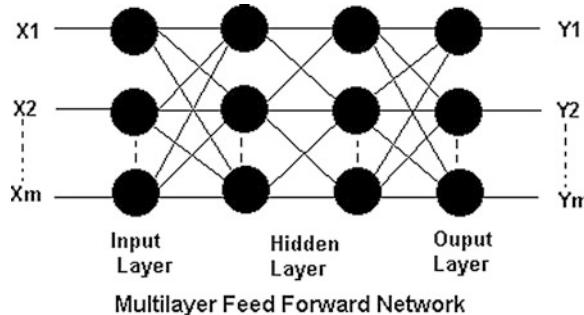


Fig. 3 The training of the network has been given by kid's handwriting and these character images contain maximum variance, which is useful for its effective training

6 Experimental Results

For analysing the impact of training samples obtained from the kids themselves, an experiment has been performed using different training samples of handwriting obtained from kids and young subjects. Similarly, testing has been carried out for these two classes, separately. For all experiments, features defined in the previous section were used. These features were given to the two hidden-layer neural network. The training was done for 260 kid's written characters and 260 for younger people (normal writing). The sample used for testing was same for both the trained classifier, i.e., for kids and young objects (normal writing). The results shown below have been achieved with effective running time and less complexity. The following four combinations were tested with the algorithm depending on the types of samples used for training and testing:

- Training sample—kids and testing samples—kids (Table 1)
- Training sample—kids and testing samples—normal (Table 2)
- Training sample—normal and testing samples—normal (Table 3)
- Training sample—normal and testing samples—kids (Table 4)

From the experimental results depicted in tables, it can be seen that the training samples from kids used, affects the performance of character recognition system. From Tables 1 and 4, it is clearly visible that kid's handwriting can be recognized more reliably when training is accomplished using kids samples as compared to that

Table 1 Recognition performance with training sample—kids and testing samples—kids

No. of training samples (kid's writing)	No. of tested sample (kid's writing)	No. of hidden layer	No. of epochs	Percentage of recognition accuracy
260	30	2	1,200	100
260	26	2	1,200	96.15
260	26	2	1,200	100

Table 2 Recognition performance with training sample—kids and testing samples—normal

No. of training samples (kid's writing)	No. of tested sample (normal writing)	No. of hidden layer	No. of epochs	Percentage of recognition accuracy
260	18	2	1,200	100
260	17	2	1,200	92.667
260	16	2	1,200	93.7,500

Table 3 Recognition performance with training sample—normal and testing samples—normal

No. of training samples (normal writing)	No. of tested sample (normal writing)	No. of hidden layer	No. of epochs	Percentage of recognition accuracy
260	18	2	1,200	87.500
260	17	2	1,200	77.778
260	16	2	1,200	87.500

Table 4 Recognition performance with training sample—normal and testing samples—kids

No. of training samples (normal writing)	No. of tested sample(kid's writing)	No. of hidden layer	No. of epochs	Percentage of recognition accuracy
260	30	2	1,200	60
260	26	2	1,200	55
260	26	2	1,200	65

with normal person handwriting. Another interesting fact that can be deduced from Tables 1 and 2 is that if kid's samples are used for training, then recognition performance for even normal person's handwriting is boosted. These results support the use of kid's handwriting for training in character recognition system in order to make system robust.

7 Conclusion

After using kid's writing samples and analyse that how human behaviour affects neural network learning, the problems of offline handwritten English character recognition has been taken, i.e., not only because of the great amount of variations in human handwriting, but also because of the overlapped and joined characters. Recognition approaches heavily depend on the nature of the data to be recognized. In this study, kid's handwriting-based handwritten character recognition system is proposed. The main emphasis in this work is in analysing the effect of training samples obtained from kids, on the recognition performance of character

identification. The positive impact of kids training set on the performance of the neural network is shown by differences in results after using kids and young objects both. The robustness of the model for the various test inputs is mainly attributed to the exhaustive and complete exploration of the data space by the neural network during the training. This study also reveals that the neural networks can be trained better for simulating the human behaviour if they are trained like human beings do.

Similarly, neural networks trained by kid's writing data sample are able to recognize kid's writing test sample inputs with an increased average recognition rate of 38.7 % as compared to neural networks trained from young subject writing data samples.

This also opens a question that whether it is possible to create an artificial brains same as of human by training an artificial neural network in similar fashion as a child is trained.

References

1. Fausett, L.: *Fundamentals of Neural Network*, Prentice Hall, Englewood cliff (1994)
2. Rojas, R.: *Neural Network*. Springer, Berlin (1996)
3. Perry, B.D.: How the brain learns better from <http://teacher.scholastic.com/professional/buceperry/brainlearns.html>
4. Zhang, Z., Vanderhaegen, F., Millot, P.: Prediction of human behaviour using artificial neural networks advances in machine learning and cybernetics. *Lect. Notes Comput. Sci.* **3930**, 770–779 (2006)
5. James, K.H., Atwood, T.P.: The role of sensorimotor learning in the perception of letter-like forms: tracking the causes of neural specialization for letters. *Cogn. Neuropsychol.* **26**(1), 91–100 (2009)
6. James, K.H., Engelhardt, L.: The effects of handwriting experience on functional brain development in pre-literate children. *Trends Neurosci. Edu.* Article in press (2013)
7. Mishra, B., Saurabh, P., Verma, B.: A novel approach to classify high dimensional datasets using supervised manifold learning. In: 4th International Conference in Recent Trends in Computer, Communication and Information Technology, OBCOM, Dec 2011
8. Ois Kampi, J.F., Poirier, F., Doignon, P.: Interaction with in-vehicle electronic systems: a complete description of a neural network approach. *Neural Process. Lett.* **19**, 109–129 (2004)
9. Alonso-Weber, J.M., Sanchis, A.: A skeletonizing reconfigurable self-organizing model: validation through text recognition. *Neural Process. Lett.* **34**, 39–58 (2011). doi:[10.1007/s11063-011-9182-0](https://doi.org/10.1007/s11063-011-9182-0) published online: 28 May 2011
10. Nasien, D., Haron, H., Yuhaniz, S.S.: Support vector machine (SVM) for english handwritten character recognition. In: Computer Engineering and Applications Second International Conference on computer science and engineering, pp. 249–252, 19–21 Mar 2010
11. Pradeep, J., Srinivasan, E., Himavathi, S.: Neural network based handwritten character recognition system without feature extraction. In: International Conference on Computer, Communication and Electrical Technology (ICCCET), pp. 40–44, Mar 2011
12. Alvaro, A.K.S., Dela Cruz, R.L.D.J., Fonseca, D.M.T., Samonte, M.J.C.: Basic handwriting instructor for kids using OCR as an Evaluator. In: International Conference on Networking and Information Technology (ICNIT), pp. 265–268, 11–12 June 2010
13. Fahmy, H.I., Develekos, G.: Application of neural network and machine learning in network design. *IEEE J. Selected Area Comm.* **15**(2), Feb 1997

14. Perwej, Y., Chaturvedi, A.: Machine recognition of handwritten characters using neural networks. *Int. J. Comput. Appl.* **14**(2), 0975–8887 Jan 2011
15. Pal, S., Mitra, J., Ghose, S., Banerjee, P.: A projection based statistical approach for handwritten character recognition. In: International Conference on Computational Intelligence and Multimedia Application, pp. 316–320 (2007)
16. Koerich, A.L., Kalva, P.R.: Unconstrained handwritten character recognition using metaclasses of characters. In: IEEE International Conference on Image Processing ICIP, vol.2, pp. II, 542–545, 11–14 Sept. 2005
17. Pal, A., Singh, D.: Handwritten character recognition using neural network. *Int. J. comput. Sci. comm.* **1**(2), 343–350, July–Dec 2010
18. Mishra, B., Saurabh, P., Verma, B.: A novel approach to predict cancer by ensembling of kernel based dimension reduction and classifier using DNA microarrays. In: IEEE International Conference on Communication Systems and Network Technologies, pp. 646–650, June 2011
19. Gonazlez, R.C., Woods, R.E., Eddins, S.L.: Digital image processing using Matlab. ISBN: 9780070702622
20. Sivanandam, S.N., Sumathi, S., Deepa, S.N.: Introduction to neural network using Matlab 6.0. ISBN: 9780070591127
21. Tier, Ø.D., Jain, A.K., Taxt, T.: Feature extraction methods for character recognition – A survey. *Pattern Recognit.* **29**(4), 641–662 (1996). Elsevier Science Ltd
22. Sonka, M., Hlavac, V., Boyle, R.: Digital image processing and computer vision, Cenage Publication. ISBN 8131505553
23. Zheng, L., He, X.: Classification techniques in pattern recognition. In: Conference Proceedings WSCG, Jan 31–Feb 4 (2005)
24. Liwicki, M., Graves, A., Bunke, H.: Neural networks for handwriting recognition studies in computational intelligence. *Comput. Intell. Paradigms Adv. Patter Classif.* 386, 5–24 (2012)
25. Sabeeenian, R.S., Vidhya, M.: Communication in computer science and information science,1. *Info. Comm. Technol.* 101, 498–500
26. Shrivastava, N.K., Saurabh, P., Verma, B.: An efficient approach parallel support vector machine for classification of diabetes dataset, *IJCA*, pp. 19–24, **36**(6), Dec 2011
27. Hanmandlu, M., Mohan, K.R.M., Chakraborty, S.: Fuzzy logic based handwritten character recognition. In: International Conference on Image Processing, pp. 42–45, vol. 3, July 2001
28. Noor, N.M., Razaz, M., Manley-Cooke, P.: Global geometry extraction for fuzzy logic based handwritten character recognition. In: The 17th International Conference on Pattern Recognition, pp. 513–516, vol. 2, Aug 2004
29. Jesan, J.P., Lauro, D.M., Human brain and neural network behaviour a comparison, Nov 2003

An Improved RC4 with Statistical Analysis on Ciphertexts

Suman Das, Hemanta Dey and Ranjan Ghosh

Abstract RC4 has proved itself as robust enough and is trusted by many organizations. A number of researchers claimed that though this stream cipher is simple, fast, easy to implement, it has some weakness and bias in its internal states. Some researchers argued that the *swap* function of RC4 in key-scheduling algorithm (KSA) and pseudo-random generation algorithm (PRGA) is the main reason of weakness. The authors of this paper eliminated the KSA and used a mathematical process to generate the internal state array(s) of RC4. Also, the PRGA has been modified to handle two S-boxes to generate two keystream bytes in one loop. Both the algorithms, original and modified, are tested with the NIST Statistical Test Suite. It has been found that the modified RC4 is giving a better randomness in the ciphertexts, hence giving a better security.

Keywords Modified RC4 · Modified KSA · Random S-box · Irreducible polynomial · NIST Statistical Test Suite

1 Introduction

RC4 is one of the most popular stream ciphers, which is simple, efficient, fast, and easy to implement. The core of the algorithm is an internal state table and a table-shuffling principle (*swap*). It contains an initialization routine and a random number generator, where the random values are selected from the internal state array and two elements are swapped for every byte reported. RC4 is designed for fast software and hardware implementation and widely used in many commercial products

S. Das (✉)

Department of Computer Science and Engineering, University of Calcutta, Kolkata, India
e-mail: ammi.suman@gmail.com

H. Dey · R. Ghosh

Institute of Radio Physics and Electronics, University of Calcutta, Kolkata, India
e-mail: hemantadey13@gmail.com

and standards. The RC4 cryptanalysis has been mainly devoted to the statistical analysis of the output sequence, or to the initialization weaknesses.

In RC4, there are two components: key-scheduling algorithm (KSA) and pseudo-random generation algorithm (PRGA). It contains a secret array S of size N (generally, 256), in which integers (0 to $N - 1$) are swapped, depending upon two index pointers i and j in a deterministic (for i) and pseudo-random (for j) way. The KSA turns an identity permutation into a random-looking permutation and the PRGA generates the keystream bytes.

There are several works on strength and weakness of RC4. Researchers argued that due to propagation of biases in the KSA, many biases are introduced in the PRGA. It shows that there is significant interest in the cryptographic community for RC4.

In this paper, the initial value of j in PRGA has been calculated from the key values, not from 0 as stated in the original algorithm. The structure of the PRGA has also been changed to handle two S-boxes. Also, the KSA has been replaced in a mathematical way of calculating and filling the S array with the multiplicative inverses of the 256 bytes (0–255) in Galois field 2^8 [GF(2^8)]. This variant of RC4 is then statistically analyzed and compared with the original RC4, following the guidelines given by National Institute of Standards and Technology (NIST), USA, in their Statistical Test Suite, which has been coded by the authors. It has been found that though RC4 itself is quite secured to use, even after so many years of its primary design, the new variant is able to prove itself more efficient (Table 1).

2 Some Previous Works

The efficiency of RC4 depends on the efficiency of KSA and PRGA. Paul and Preneel [1] proposed to introduce more random variables in the PRGA to reduce the correlation between the internal and the external states. They described a new statistical bias in the distribution of the first two output bytes of RC4 and also

Table 1 The RC4 stream cipher

KSA	PRGA
Input: Secret Key K	Input: S-Box S – The o/p of KSA
for $i = 0, \dots, N - 1$	
$S[i] = i;$	$i = 0; j = 0;$
next i	while $TRUE$
$j = 0;$	$\{ i = i + 1$
for $i = 0, \dots, N - 1$	$j = j + S[i]$
$\{ j = j + S[i] + K[i]$	$\text{swap}(S[i], S[j]);$
$\text{swap}(S[i], S[j]);$	$z = S[S[i] + S[j]]; \}$
}	
next i	
<u>Output: S-Box S generated by K</u>	<u>Output: Random Stream Z</u>

proposed a new algorithm, namely RC4A, with much less operations per output byte and recommended to drop at least $2N$ initial bytes, where N is the size of the internal S-box.

Maitra and Paul [2] revolved the non-uniformity in KSA and proposed for additional layers over the KSA and the PRGA. They named the modified cipher as RC4+, which avoids existing weaknesses of RC4. To remove weaknesses of KSA, a three-layer architecture is presented in a scrambling phase after the initialization (KSA+). They also introduced some extra phases to improve the PRGA (PRGA+).

Roos [3] identified several classes of weak keys for RC4, showing some weakness in KSA with some important technical results. He defined strong correlation between the secret key bytes and the final keystream generated.

Sen Gupta et al. [4] thoroughly studied the RC4 designing problem from the view point of throughput. They implemented a hardware architecture to generate two keystream bytes per clock cycle using the idea of loop unrolling and hardware pipelining.

Nawaz et al. [5] introduced a new 32-bit RC4 like faster keystream generator. It has a huge internal state and offers higher resistance against state recovery attacks. This is suitable for high-speed software encryption.

Akgün et al. [6] proposed a framework, which significantly increases the success rate of key retrieval attack. They showed that KSA leaks information about the secret key, if the initial state table is known. They also detected a new bias in the KSA and proposed a new algorithm to retrieve the RC4 key in a faster way.

Tomasevic and Bojanic [7] used a strategy to favor more promising values that should be assigned to unknown entries in the RC4 table. An abstraction in the form of general conditions about the current state of RC4 has been introduced. They proposed a new technique to improve cryptanalytic attack on RC4, which is based on new information from the tree representation of RC4.

Church [8] gave the complete list of irreducible polynomials for prime moduli (2, 3, 5, 7, and 11) of each degree. The determination of the exponents provided in itself a very satisfactory control of the irreducibility of the polynomials.

Daemen and Rijmen [9] showed the process of calculating multiplicative inverse of a byte (as an 8-bit binary string) in $GF(2^8)$ with an irreducible polynomial as the modulus. They defined a process of creating cryptographic S-boxes by a mathematical process in $GF(2^8)$ field.

Das et al. [10] eliminated the swap function of KSA by using a mathematical process to fill up the internal state array of RC4, which has been found giving a better security after statistical analysis.

3 A Modified RC4

Roos [3] and others strongly discussed about the weakness of KSA and weak keys in RC4. Roos showed that there is a high probability of about 37 % for an element not to be swapped at all. He argued that in KSA, only the line of swap directly affects the

state table S while exchanging two elements, and hence, the previous line $j = j + S[i] + K[i]$ is responsible for calculating the indexes, where i is deterministic and j is pseudo-random. Therefore, the swap between two elements may happen once, more than once, or may not happen at all—thus brings a weakness in the KSA.

In this paper, the authors propose to introduce a modified RC4, where all the cells of the initial internal state array S will be filled up by calculating the multiplicative inverses of the 256 bytes (0–255) by using any one of the 30 irreducible polynomials [8] in $GF(2^8)$ as the modulus [11, 12]. In this way, the state array S can be filled up by 256 random values without using any swap function. A list of these irreducible polynomials, given in hexadecimal notations, is as follows [13, 14]:

11B	11D	12B	12D	139	13F	14D	15F	163	165	169	171	177	17B	187
18B	18D	19F	1A3	1A9	1B1	1BD	1C3	1CF	1D7	1DD	1E7	1F3	1F5	1F9

Furthermore, two S-boxes have been generated by using any two of the above-mentioned irreducible polynomials as moduli and the PRGA has been modified to produce two keystream bytes, one from each S-box, in a single-loop cycle. The modified PRGA is as follows:

```


$$j_1 = (j_1 + S_1[i] + S_2[i_2]);$$


$$swap(S_1[i], S_1[j_1]);$$


$$z = S_1[j_1]);$$


$$i++;$$


$$\longrightarrow$$


$$j_2 = (j_2 + S_2[i] + S_1[i_1]);$$


$$swap(S_2[i], S_2[j_2]);$$


$$z = S_2[j_2]);$$


$$i++;$$


```

The initial values of j_1 and j_2 has been calculated from the key values, not from 0 as stated in the original algorithm or in other variants of RC4, to give more dynamic values to these variables as the initial starting points of the PRGA. The basic aim of the authors is to make the outputs of the algorithm more random to make it more difficult to the intruder to break it.

The outputs of RC4 (original and modified) have been tested statistically using the guidance of NIST, by the NIST Statistical Test Suite. For both the algorithms, a same text file has been encrypted 500 times by using 500 same encryption keys, generating 500 ciphertexts for each algorithm, each of which contains at least 1,342,500 bits, as recommended by NIST. The two sets of ciphertexts are then tested statistically to find out if security varies for the original and the modified algorithms.

4 The NIST Statistical Test Suite

NIST developed a Statistical Test Suite, which is an excellent and exhaustive document consisting of 15 tests developed to test various aspects of randomness in binary sequences produced by cryptographic algorithms [15, 16]. The tests are as follows:

1. *Frequency (Monobit) Test*: Number of 1's and 0's in a sequence should be approximately the same, i.e., with probability $1/2$.
2. *Frequency Test within a Block*: Whether frequency of 1's in an M -bit block is approximately $M/2$.
3. *Runs Test*: Whether number of runs of 1's and 0's of various lengths is as expected for a random sequence.
4. *Test for Longest-Run-of-Ones in a Block*: Whether the length of the longest run of 1's within the tested sequence (M -bit blocks) is consistent with the length of the longest run of 1's as expected.
5. *Binary Matrix Rank Test*: Checks for linear dependence among fixed length substrings of the sequence, by finding the rank of disjoint submatrices of it.
6. *Discrete Fourier Transform Test*: Detects periodic features in the sequence by focusing on the peak heights in the DFT of the sequence.
7. *Non-overlapping Template Matching Test*: Occurrences of a non-periodic pattern in a sequence, using a non-overlapping m -bit sliding window.
8. *Overlapping Template Matching Test*: Occurrences of a non-periodic pattern in a sequence, using an overlapping m -bit sliding window.
9. *Maurer's Universal Statistical Test*: Whether or not the sequence can be significantly compressed without loss of information, by focusing on the number of bits between matching patterns.
10. *Linear Complexity Test*: Finds the length of a linear feedback shift register (LFSR) to generate the sequence—longer LFSRs imply better randomness.
11. *Serial Test*: Determines number of occurrences of the 2^m m -bit overlapping patterns across the entire sequence to find uniformity—every pattern has the same chance of appearing as of others.
12. *Approximate Entropy Test*: Compares the frequency of all possible overlapping blocks of two consecutive/adjacent lengths (m and $m + 1$).
13. *Cumulative Sums Test*: Finds if the cumulative sum of a sequence is too large or small and focuses on maximal excursion of random walks, which should be near 0.
14. *Random Excursions Test*: Finds if number of visits to a state within a cycle deviates from expected value, calculates the no. of cycles having exactly K visits in a cumulative sum random walk.
15. *Random Excursions Variant Test*: Deviations from the expected visits to various states in the random walk, calculates the number of times that a state is visited in a cumulative sum random walk.

In each test, for a bit sequence, NIST adopted different procedures to calculate the P values (probability values) for different tests from the observed and expected values under the assumption of randomness. The Test Suite has been coded by the authors and used to study the randomness features of RC4 and the modified RC4.

5 Results and Discussions

After analyzing the outputs of the original RC4 and modified RC4, using the NIST Statistical Test Suite, as described above, it has been found that the modified algorithm creates a tweak in RC4 to increase its security. The final analysis and comparison is displayed in Table 2, where the POP (Proportion of Passing) status and Uniformity Distribution of NIST tests for these two algorithms are displayed and compared. The best values of a particular test for each algorithm are shaded (in rows), and then, the numbers of shaded cells for each are counted (in columns). The higher count (here for the modified RC4) gives a better result for that particular algorithm, which shows that this one has a better security than the other, at least for this particular data set.

POPs and Uniformity Distributions generated by these 2 algorithms for the 15 tests, compared to the expected values [15], are displayed in Table 3a and b. Distributions of P values generated by the algorithms, from which the Uniformity Distribution has been calculated [15], are displayed in Table 4a and b. Here, the interval between 0 and 1 is divided into 10 subintervals, and the P values that lie within each subinterval are counted and displayed. These P values should be uniformly distributed in each subinterval [15]. Histograms on distribution of P values for two tests (4 and 8) are displayed in Figs. 1a, b and 2a, b, respectively.

Scattered graphs on the POP status for the 15 tests are displayed in Fig. 3a and b, which examine the proportion of sequences that pass a test. A threshold value (expected POP) has been calculated following the guidance given by NIST. If the

Table 2 Comparison of POP status and Uniformity Distribution generated by the 15 NIST tests for RC4 and the modified RC4

Test↓	POP Status		Uniformity Distribution	
	Modified RC4	RC4	Modified RC4	RC4
1	0.984000	0.988000	3.314076 ⁻⁰¹	4.154218 ⁻⁰¹
2	0.992000	0.992000	5.301203 ⁻⁰¹	4.904834 ⁻⁰¹
3	0.988000	0.992000	9.540148 ⁻⁰¹	8.920363 ⁻⁰¹
4	0.988000	0.982000	4.673217 ⁻⁰¹	5.790211 ⁻⁰¹
5	0.990000	0.984000	5.101528 ⁻⁰¹	2.492839 ⁻⁰¹
6	0.986000	0.980000	5.328562 ⁻⁰²	4.170881 ⁻⁰²
7	0.992000	0.990000	5.666881 ⁻⁰¹	8.272794 ⁻⁰¹
8	0.988000	0.992000	9.502466 ⁻⁰¹	2.224804 ⁻⁰¹
9	0.986000	0.982000	1.835471 ⁻⁰¹	3.856456 ⁻⁰²
10	0.990000	0.992000	1.855553 ⁻⁰¹	5.462832 ⁻⁰¹
11	0.990000	0.982000	6.163052 ⁻⁰¹	1.699807 ⁻⁰¹
12	0.990000	0.992000	3.756609 ⁻⁰²	2.953907 ⁻⁰¹
13	0.991000	0.995000	3.586413 ⁻⁰¹	8.201435 ⁻⁰¹
14	0.988500	0.983500	7.858791 ⁻⁰¹	6.729885 ⁻⁰¹
15	0.990333	0.985889	1.673032 ⁻⁰²	8.386675 ⁻⁰²
Total:	9	7	8	7

Table 3 (a) POP status and Uniformity Distribution generated for the modified RC4, (b) POP status and Uniformity Distribution generated for RC4

Test↓	Expected POP	Observed POP	Status	Uniformity Distribution	Status
(a)					
1	0.976651	0.984000	Successful	3.314076^{-01}	Uniform
2	0.976651	0.992000	Successful	5.301203^{-01}	Uniform
3	0.976651	0.988000	Successful	9.540148^{-01}	Uniform
4	0.976651	0.988000	Successful	4.673217^{-01}	Uniform
5	0.976651	0.990000	Successful	5.101528^{-01}	Uniform
6	0.976651	0.986000	Successful	5.328562^{-02}	Uniform
7	0.976651	0.992000	Successful	5.666881^{-01}	Uniform
8	0.976651	0.988000	Successful	9.502466^{-01}	Uniform
9	0.976651	0.986000	Successful	1.835471^{-01}	Uniform
10	0.976651	0.990000	Successful	1.855553^{-01}	Uniform
11	0.980561	0.990000	Successful	6.163052^{-01}	Uniform
12	0.976651	0.990000	Successful	3.756609^{-02}	Uniform
13	0.980561	0.991000	Successful	3.586413^{-01}	Uniform
14	0.985280	0.988500	Successful	7.858791^{-01}	Uniform
15	0.986854	0.990333	Successful	1.673032^{-02}	Uniform
(b)					
1	0.976651	0.988000	Successful	4.154218^{-01}	Uniform
2	0.976651	0.992000	Successful	4.904834^{-01}	Uniform
3	0.976651	0.992000	Successful	8.920363^{-01}	Uniform
4	0.976651	0.982000	Successful	5.790211^{-01}	Uniform
5	0.976651	0.984000	Successful	2.492839^{-01}	Uniform
6	0.976651	0.980000	Successful	4.170881^{-02}	Uniform
7	0.976651	0.990000	Successful	8.272794^{-01}	Uniform
8	0.976651	0.992000	Successful	2.224804^{-01}	Uniform
9	0.976651	0.982000	Successful	3.856456^{-02}	Uniform
10	0.976651	0.992000	Successful	5.462832^{-01}	Uniform
11	0.980561	0.982000	Successful	1.699807^{-01}	Uniform
12	0.976651	0.992000	Successful	2.953907^{-01}	Uniform
13	0.980561	0.995000	Successful	8.201435^{-01}	Uniform
14	0.985280	0.983500	Unsuccessful	6.729885^{-01}	Uniform
15	0.986854	0.985889	Unsuccessful	8.386675^{-02}	Uniform

proportion falls outside of (i.e., less than) this expected value, then there is evidence that the data are not random [15]. If most of the values are greater than this expected value, then the data are considered to be random. For a particular algorithm, the more number of POPs tend to 1 for the 15 tests, the more random, i.e., more secured will be the data sequence.

Table 4 (a) Distribution of P values generated for the modified RC4, (b) distribution of P values generated for RC4

Test↓	1	2	3	4	5	6	7	8	9	10
(a)										
1	51	51	47	58	66	42	41	50	50	44
2	57	50	53	42	57	43	40	50	49	59
3	46	51	54	53	46	56	54	46	50	44
4	51	52	49	52	55	44	34	49	59	55
5	61	50	48	49	41	40	57	46	56	52
6	56	56	50	41	41	44	64	65	42	41
7	43	53	47	51	42	45	46	53	59	61
8	55	45	44	51	50	57	51	48	46	53
9	64	60	56	50	45	46	48	53	39	39
10	57	43	58	48	41	34	50	57	55	57
11	95	113	97	104	108	107	93	95	105	83
12	47	56	44	64	58	53	37	45	61	35
13	105	99	100	87	90	95	96	123	95	110
14	412	370	393	378	416	412	395	408	411	405
15	815	853	920	917	937	890	898	970	929	871
(b)										
1	69	49	49	48	53	47	45	46	50	44
2	50	48	54	38	52	44	63	47	48	56
3	56	58	49	43	55	47	49	48	50	45
4	39	60	51	50	47	59	52	43	52	47
5	59	55	49	49	48	38	45	65	50	42
6	57	46	47	44	60	37	48	69	39	53
7	42	44	47	48	53	51	57	49	50	59
8	47	56	41	56	43	49	65	57	41	45
9	61	54	57	54	58	53	33	41	54	35
10	52	43	57	47	44	45	64	46	53	49
11	100	114	123	103	99	103	83	97	89	89
12	50	56	66	49	55	50	40	49	41	44
13	91	105	106	97	100	103	103	111	99	85
14	439	399	415	381	398	383	400	384	402	399
15	956	853	902	890	949	902	896	934	833	885

Horizontal ranges for a and b: 1 0.0–0.1, 2 > 0.1–0.2, 3 > 0.2–0.3, ..., 10 > 0.9–1

Finally, it has been observed that besides using the original KSA, a suitable mathematical process can be used to generate a secured internal state array for RC4, and also, a modified PRGA with two output bytes in a single loop may give better randomization in the ciphertexts of RC4.

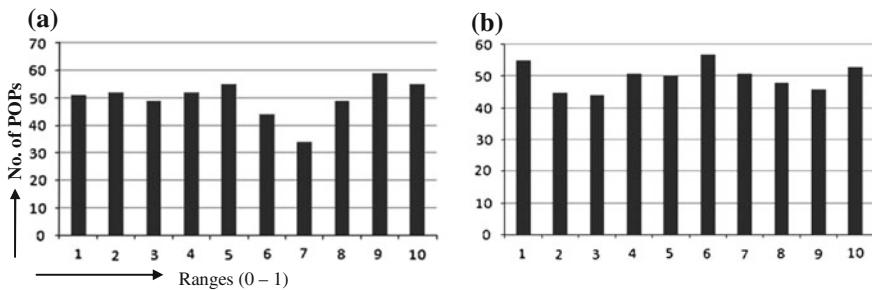


Fig. 1 a and b Histograms for P value distribution of test 4 and 8 for modified RC4

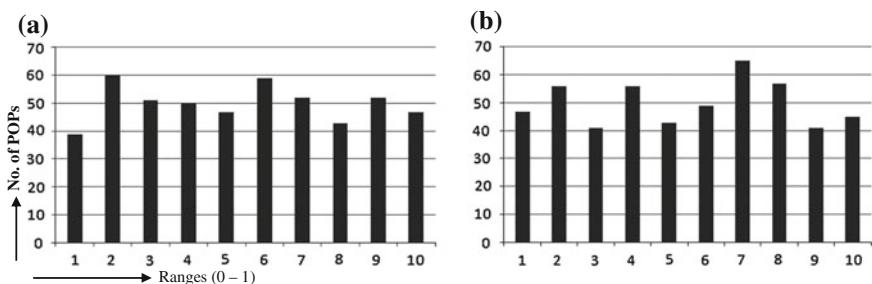


Fig. 2 a and b Histograms for P value distribution of test 4 and 8 for RC4

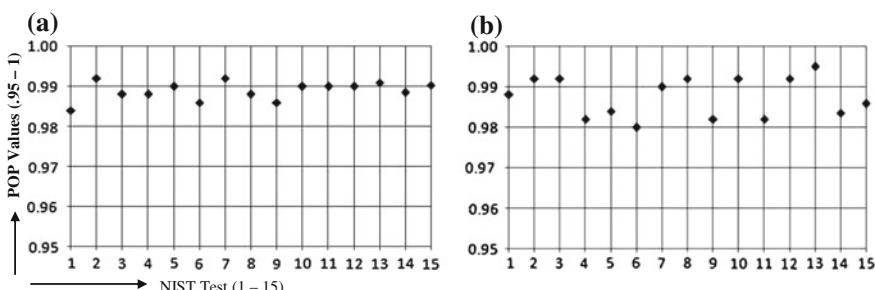


Fig. 3 a and b Scattered graph on POP status on 15 NIST tests for modified RC4 and RC4

6 Conclusion

Though RC4 is robust enough, but one can envisage that the internal state array is the convenience available to the intruders due its weakness and some internal biases, as argued by a number of researchers. Undoubtedly, the security of RC4 will be enhanced if the said state array(s) can be replaced by using some mathematical or randomization techniques where they become stronger for different applications.

In this paper, randomness has been introduced in respect of forming the state array (s), with a modified intelligent PRGA, based on which the security of RC4 has been enhanced.

The modified RC4 is found to stand in the better merit list comparing to the standard RC4. It seems that security in RC4 will be enhanced by driving a mathematical process to generate the initial internal state array, accompanied by an improved PRGA. Also, the user can choose and generate any random state array S by using his/her own choice of modulus from a large set of options. In the case of suspicion of a trapdoor in the ciphertext, a state array might be replaced by another one by the user. Other mathematical procedures/models and modifications may also be used to modify the RC4 algorithm and studies on them are required to find better opportunities to generate more secured RC4 keystreams.

References

1. Paul, S., Preneel, B.: A new weakness in the RC4 keystream generator and an approach to improve the security of the cipher. In: FSE 2004, LNCS, vol. 3017, pp. 245–259. Springer, Heidelberg (2004). <http://www.iacr.org/archive/fse2004/30170244/30170244.pdf>. Last accessed on 2 July 2014
2. Maitra, S., Paul, G.: Analysis of RC4 and proposal of additional layers for better security margin. In: INDOCRYPT, Lecture Notes in Computer Science, vol. 5365, pp. 40–52. Springer, Berlin (2008). <http://eprint.iacr.org/2008/396.pdf>. Last accessed on 2 July 2014
3. Roos, A.: A Class of Weak Keys in the RC4 Stream Cipher. Post in sci.crypt (1995)
4. Sen Gupta, S., Chattopadhyay, A., Sinha, K., Maitra, S., Sinha, B.P.: High-performance hardware implementation for RC4 stream cipher. IEEE Trans. Comput. **82**(4) (2013). (Last accessed on 2 July 2014)
5. Nawaz, Y., Gupta, K.C., Gong, G.: A 32-bit RC4-like keystream generator, IACR Eprint archive, 2005. <https://eprint.iacr.org/2005/175.pdf>. Last accessed on 2 July 2014
6. Akgün, M., Kavak, P., Demirci, H.: New results on the key scheduling algorithm of RC4. In: INDOCRYPT, Lecture Notes in Computer Science, vol. 5365, pp. 40–52. Springer, Berlin (2008). http://link.springer.com/content/pdf/10.1007/978-3-540-9754-5_4.pdf
7. Tomašević, V., Bojanić, S.: Reducing the state space of RC4 stream cipher. In: Bubak, M., et al. (eds.) ICCS 2004, LNCS, vol. 3036, pp. 644–647. Springer, Berlin, Heidelberg (2004). http://link.springer.com/chapter/10.1007%2F978-3-540-24685-5_110#page-1. Last accessed on 2 July 2014
8. Church, R.: Tables of irreducible polynomials for first four prime moduli. Ann. Maths. 2nd Sr. 36(1), 198–209 (1935). <http://www.jstor.org/stable/1968675>
9. Daemen, J., Rijmen, V.: AES proposal: Rijndael, version 2, Submitted to NIST, Mar 1999. <http://csrc.nist.gov/encryption/aes>
10. Das, S., Dey, H., Ghosh, R.: Comparative study of randomness of RC4 and a modified RC4. In: International Congress on Electronics Engineering and Computer Science, IEMCONG-2014, Kolkata, India (2014)
11. Foruzan, B.: Cryptography and Network Security. Tata McGraw-Hill, New Delhi (2007). Special Indian Edition
12. Stinson, D.R.: Cryptography—Theory and Practice. Department of Combinatorics and Optimization, University of Waterloo, Ontario (2002)
13. FIPS: Announcing AES, 2001. <http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf>

14. FIPS, PUB 197: The official AES standard, 2001-11-26. Retrieved 29 Apr 2010. <http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf>
15. National Institute of Standard & Technology (NIST), Technology Administration, U.S. Department of Commerce, A statistical test suite for RNGs & PRNGs for cryptographic applications, 2010. <http://csrc.nist.gov/publications/nistpubs800/22rec1SP800-22red1.pdf>
16. Kim, S.J., Umeno, K., Hasegawa, A.: Corrections of the NIST statistical test suite for randomness, Communications Research Lab., Inc. Admin. Agency, Tokyo, Japan (2004)

Gathering of Swarm of Mobile Robots in Presence of Horizontal Obstacles

U. Baveenther and Deepanwita Das

Abstract This paper presents a distributed algorithm for gathering of swarm mobile robots in the presence of horizontal line obstacles. A swarm of mobile robots with unlimited visibility are randomly located within bounded area. Each robot performs a deterministic algorithm, and as a result, they will finally gather at some point without exactly reaching that point. We have followed *asynchronous* timing model and *full-compass* models.

Keywords Gathering • Swarm robots • Horizontal obstacles

1 Introduction

Gathering problem is very important problem in the field of swarm robots. The main objective of swarm robotics research is to observe the behaviors of a group of simple autonomous robots in various environments and to coordinate large groups of simple and identical robots by using local rules. Beni [1] describes this as follows:

The group of robots is not just a group. It has some special characteristics, which are found in swarms of insects, that is, decentralized control, lack of synchronization, simple and (quasi) identical members.

Based on the swarm behavior, different problems were discussed in the works related to swarm robotics such as *geometric pattern formation*, *flocking*, *coverage*, *partitioning*, *spreading*, *searching*, *gathering*, and *convergence*. Among these problems, we have focused our research on the area related to the *gathering* problem. *Gathering* is defined as a process in which all the mobile robots within a bounded region come close to each other and gather toward a point within finite amount of time, without reaching that point exactly. Initially, all the robots are

U. Baveenther (✉) · D. Das
National Institute of Technology, Durgapur, Mahatma Gandhi Avenue,
Durgapur 713209, West Bengal, India
e-mail: baveenther@gmail.com

scattered within bounded area. Each and every robot may or may not be visible to the others. This proposed algorithm results in gathering of all robots at some point without reaching that point exactly. The basic idea of our approach is that each and every robot searches other robots by moving around the obstacles nearest to it. If it finds any robot, then they will gather and form a group and together this group will search for other robots in the system. Finally, the algorithm will complete the gathering task within finite amount of time.

A large amount of research works have been done in this area. Cieliebak et al. [2, 3] proposed an algorithm where initially all the robots are visible to each other. Based on the initial configuration of robots and total number of robots in the system, a center of gravity method is used to gather all the mobile robots. For example, if the total number of robots in the system is 2, i.e., $N = 2$, then they will form a line. Each robot calculates a center of the line. Finally, both of them will gather at midpoint of the line. Similarly, if $N = 3$, the robots will form a triangle. Then, each of them calculates the center of the pattern and will move to that point. Though the robots are able to gather at a single point, the algorithm is not so efficient because the whole environment is known to all the robots as all the robots are visible to each other throughout the process.

Agmon and peleg [4] proposed a fault-tolerant algorithm for a given task. The algorithm ensures that in a system of N robots, if at most f robots fail in any execution, the task is achieved by all nonfaulty robots, regardless of the actions taken by the faulty ones. In the gathering task, when faults are introduced into the system, the requirement applies only for the nonfaulty robots, i.e., if f robots fail, then all the remaining $N - f$ nonfaulty robots are required to occupy a single point within a finite time. The algorithm only emphasizes a fault-tolerant system but not considered the occurrence of faults due to the presence of obstacles.

Flochinni et al. [5] present an algorithm that allows anonymous, oblivious robots with limited visibility to gather at a single point within a finite amount of time. It has also used the concept of maintaining mutual visibility graph throughout the process. The algorithm works well maintaining a mutual visibility graph but requires great amount of computational powers in order to perform better. Also, maintaining the mutually visible graph throughout the process makes the robot a complex one.

Most of the literature studies mainly concentrate on the gathering problems which consider a priori known environment, without considering the presence of obstacles. Moreover, maintaining graphs such as mutual graph makes the solution unnecessarily tough. In this paper, we have presented an algorithm that considers a similar problem of gathering with a stronger model than the literature studied. We have considered the presence of horizontal obstacles. The internal environment is completely unknown to the robots. The robots have unlimited visibility, though their view may get restricted due to the presence of obstacles, adding a realistic feature to the whole process. The algorithm guarantees collision-free gathering within finite amount of time.

2 Problem Definition, Characteristics, Models, and Assumptions

This article is aiming to solve a traditional gathering problem of the swarm of mobile robots in the presence of horizontal obstacles. The concept of *Leader-Follower* is used where one robot acts as *Leader* and remaining robots act as *Followers* who will follow the instructions communicated by the *leader*.

2.1 Characteristics

The robots are assumed to have the following characteristics:

1. **Identical:** All the robots are identical in all respect. All the robots are assumed to be point robots.
2. **Mobile:** All robots are allowed to move on a plane.
3. **Visibility:** The robots are assumed to have unlimited visibility. The robots' visibility may get restricted in the presence of any obstacle, whereas limited visibility and restricted visibility are two different facts.
4. **Communication:** Robots are allowed to communicate with each other when they are visible to each other.
5. **Nonoblivious:** Robots will have some amount of memory and can retain the information gathered in the previous computational cycles.

We have assumed that the robots observe and store certain parameters (Table 1).

2.2 Models and Assumptions

We have assumed *synchronous* timing model and *full-compass* model for direction and orientation. The models are as follows:

Asynchronous Model: The robots operate on independent computational cycles of variable lengths. They do not share any common clock.

Full compass: Directions of both axes and the positive orientation of the axes are common to all the robots. However, it is assumed that the direction of the positive y -axis is 90° counterclockwise to the positive direction of the x -axis. Each robot has its local coordinate system in which it is located at the origin.

Table 1 Parameters to be considered by the robots together with their initial values

Parameters	Descriptions	Initial value
Flag	To find leader or follower	1
nvr	Number of visible robots	0
nf	Total number of followers	0
(l_x, l_y)	Leader robot coordinate	$(0, 0)$

This paper considers the following assumptions:

1. The obstacles are assumed to be horizontal obstacles with fixed length l .
2. No two robots can occupy the same position.
3. The visibility range of the robot is infinite.
4. Initially, all the robots will act as *Leader*.
5. The *gathering* operation is assumed to be an *Atomic operation*. Once it will start, the gathering will be completed without any interruptions.

3 Proposed Algorithm

The gathering problem addressed here has to be solved by five algorithms as follows.

1. **Initialization algorithm:** Initially, all the robots in the system execute this algorithm. The parameter initialization of robot will be done by this algorithm.
Input: All the robots in the system.
Output: All the parameters are initialized.
2. **Pathfinder algorithm:** This algorithm is used to calculate the hexagon points around obstacle. Then, the leader robot will move to this path for searching other robots.
Input: Obstacle coordinate chosen by robot.
Output: Hexagon points around obstacle.
3. **Gathering algorithm:** This algorithm results in the gathering of all robots.
Input: Hexagon points around obstacle.
Output: All the robots gathered in a location.
4. **Leader–Follower algorithm:** This algorithm is used to find a leader and its followers based on the location coordinates of neighboring robots.
Input: Location coordinates of the visible robots.
Output: One leader and one/many followers in a group.
5. **Follower algorithm:** Only the followers and semi-leaders will execute this algorithm. It is used to convert the position coordinate that was sent by leader w. r.t. followers' local coordinate system.
Input: Coordinate sent by leader.
Output: New coordinate w.r.t. its local coordinate system.

Algorithm 1 (*initrobot*)

Step 1: A robot R observes the area and stores the coordinates of all the visible robots and obstacles in two separate arrays. The robots' coordinates will be stored in $\text{Robot}_{\text{Array}}$. Coordinates of the obstacles and their distance from the robot R will be stored in $\text{Obs}_{\text{Array}}$. R will also store the total number of visible robots in a variable nvr .

Step 2: If the robot R is not able to view any robot(s), i.e., the $\text{Robot}_{\text{Array}}$ is blank, then it will start executing the next algorithm skipping the remaining

steps. Otherwise, R will execute the *leader-follower* algorithm to identify itself as a leader or a follower.

Step 3: If R is identified as a *leader*, then it will execute the next algorithm skipping the remaining steps. If R is identified as a *follower*, then it will execute the *follower* algorithm.

Algorithm 2 (*pathfinder*)

Step 1: Robot R selects an obstacle having minimum distance from itself from $\text{Obs}_{\text{Array}}$. If that obstacle is not visited yet by R , i.e., its coordinate is not in Vis_{Obs} , then it will go to the next step. Otherwise, R will select next minimum distance obstacle from $\text{Obs}_{\text{Array}}$, then again it executes this step.

Step 2: If both *left node* and *right node* of that obstacle are visible to robot R as shown in Fig. 1, R will execute the next step. Otherwise, R will move itself in a position where both nodes of that obstacle are visible and execute algorithm *nitrobot*.

Step 3: R stores the coordinates of the selected obstacle in Vis_{Obs} array. Next, it selects few equal points located vertically above and below along with the *left node* and *right node* of that obstacle. R will calculate four midpoints called as *topleft-center*, *topright-center*, *bottomright-center*, and *bottomleft-center* using the center point formula $(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2})$. For example, see Fig. 2a. The *left node* and *right node* of obstacle coordinates are $(-2, 3)$ and $(3, 3)$ w.r.t. robot R . Assumed points above from *left node* are $(-2, 4), (-2, 5), (-2, 6), (-2, 7)$ and below from *left node* are $(-2, 2), (-2, 1), (-2, 0), (-2, -1)$. Similarly, assumed points above and below from *right node* are $(3, 4), (3, 5), (3, 6), (3, 7)$ and $(3, 2), (3, 1), (3, 0), (3, -1)$. Then, midpoint calculations are as follows.

$$\text{topleft-center} = ((-2 + -2)/2, (3 + 7)/2) = (-2, 5)$$

$$\text{topright-center} = ((3 + 3)/2, (3 + 7)/2) = (3, 5)$$

$$\text{bottomright-center} = ((3 + 3)/2, (3 + -1)/2) = (3, 1)$$

$$\text{bottomleft-center} = ((-2 + -2)/2, (3 + -1)/2) = (-2, 1)$$

Finally, the four midpoints are $(-2, 5), (3, 5), (3, 1)$, and $(-2, 1)$.

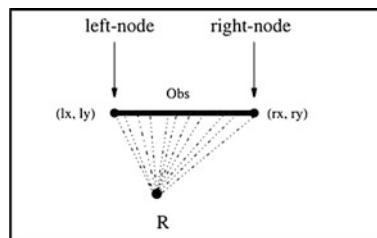


Fig. 1 Left node and right node of the obstacle (Obs) are visible to robot R

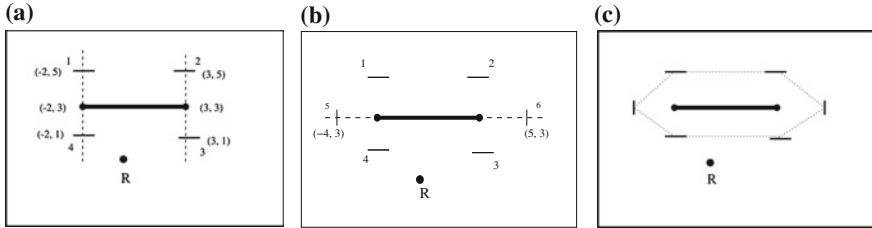


Fig. 2 Obstacle with four midpoints. 1 topleft-center, 2 topright-center, 3 bottomright-center, and 4 bottomleft-center

Step 4: R will select few points, located horizontally to the left side of the *left node* and right side of the *right node*. It then computes two midpoints from them as *left-center* and *right-center* such that all the 6 midpoints form a hexagon around the obstacle (see in Fig. 2c). From Fig. 2b, assumed left-side points from *left node* are $(-3, 3)$, $(-4, 3)$, $(-5, 3)$, and $(-6, 3)$. Similarly, right-side points from *right node* are $(4, 3)$, $(5, 3)$, $(6, 3)$, and $(7, 3)$. Midpoint calculations are

$$\text{left-center} = ((-2 + -6)/2, (3 + 3)/2) = (-4, 3)$$

$$\text{right-center} = ((3 + 7)/2, (3 + 3)/2) = (5, 3)$$

Then, two midpoints are $(-4, 3)$ and $(5, 3)$.

Step 5: The robot R will check whether there is any obstacle from itself to each and every center point as shown in Fig. 3. If it finds any obstacle in this way, it will avoid this point to be considered as a starting center point. Otherwise, it will calculate the distance from itself to that center point. Finally, it will select a starting center point which is having minimum distance from its position. Then, the robot R will store all the center points from its starting center point in array called $\text{PathCoord}_{\text{Array}}$, which forms the circle from the starting center point in clockwise direction. Then, it will start executing the next algorithm.

For example, as shown in Fig. 3, starting center point of robot R is *bottomleft-center*, i.e., $m1$ and its $\text{PathCoord}_{\text{Array}}$ will be coordinates of *bottomleft-center* ($m1$), *left-center* ($m2$), *topleft-center* ($m3$), *topright-center* ($m4$), *right-center* ($m5$), *bottom-right-center* ($m6$).

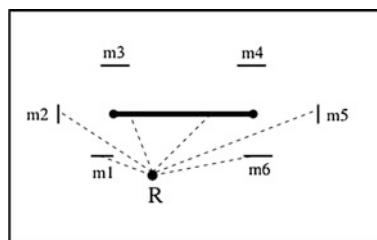


Fig. 3 Robot R moving cycle: $m1 \rightarrow m2 \rightarrow m3 \rightarrow m4 \rightarrow m5 \rightarrow m6 \rightarrow m1$

Algorithm 3 (*gathering*)

Step 1: The leader robot R will move to the first location extracted from $\text{PathCoord}_{\text{Array}}$. Upon reaching, the leader robot will send this current position to its follower(s) if it has any. After receiving the instruction from the leader, follower robot(s) will execute the *follower* algorithm.

Step 2: Before moving to the next location in the $\text{PathCoord}_{\text{Array}}$, the leader robot R observes and stores the location of the newly viewed robot(s) in $\text{Robot}_{\text{Array}}$ and will execute the *leader-follower* algorithm.

Step 3: Robot R will check whether there is any obstacle from itself to next center point. If it finds any obstacle, then all the visible obstacles from its position and their corresponding distances from itself will be stored in an array called $\text{Obs}_{\text{Array}}$. Then, R will execute the *pathfinder* algorithm. Otherwise, it will move to the next center point. After reaching this point, the leader robot will send its position to its follower(s).

Step 4: The leader robot R will continue from step 2 to step 3 until it finishes traversing one cycle that consists of all the points stored in $\text{PathCoord}_{\text{Array}}$.

Step 5: After one cycle is finished, the leader robot R finds a new obstacle. Again, it will store the coordinates and distances in $\text{Obs}_{\text{Array}}$ and will execute the *pathfinder* algorithm. Once all the obstacles in the system are visited by robot R , then the whole process will be terminated.

Algorithm 4 (*leader-follower*)

Step 1: The robot R broadcasts all the coordinates in $\text{Robot}_{\text{Array}}$ along with its own flag and nvr. If the robot can view m number of robots in any instance and their corresponding coordinates are $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$, then robot R will broadcast $(\text{flag}, \text{nvr}, (x_1, y_1), (x_2, y_2), \dots, (x_m, y_m))$ to all other visible robots. Here, $x_1 > x_2 > \dots > x_m$ and $x_1 > 0$. In the similar manner, other robots also broadcast their gathered information. For example, as shown in Fig. 4, initial value of the $\text{Robot}_{\text{Array}}$ for robot R is $S(1, 4), P(3, 4)$. So the robot R broadcasts $(1, 2, S(1, 4), P(3, 4))$ to robots S and P . Here, $\text{flag} = 1$ and $\text{nvr} = 2$.

Step 2: After receiving coordinates from all of its neighbor, the robot R will check the received flag values one by one. If the flag is a leader, then R will convert the coordinates of the corresponding robot w.r.t its own coordinate system. After the conversion, there may be two possible cases:

Case I: If results are $(0, 0)$, robot R will neglect this as it signifies its own local position.

Case II: If results are other than $(0, 0)$ and if the coordinate is already in $\text{Robot}_{\text{Array}}$, then R will neglect that. Otherwise, R will store this newly converted coordinate as a newly viewed robot coordinate and update the value of nvr.

For example, as shown in Fig. 4, robot R received the coordinates $(1, 3, R(-3, -4), S(-2, 0), M(6, -4))$ from robot P . Robot R already knows the coordinate of robot P . Robot P coordinate w.r.t. robot R is $(3, 4)$. By adding the first value $(-3, -4)$ with $(3, 4)$, it will get $(0, 0)$. Then, the robot R will neglect this value.

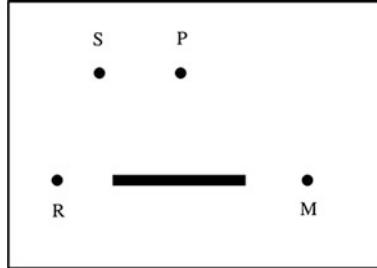


Fig. 4 Initial positions of robots R , S , P , and M . Here, the robot R can be able to see robots S and P but not M , because of the obstacle. Initially, all the robots are leader. W.r.t. robot R , the visible robot coordinates are $S(1, 4)$ and $P(3, 4)$. Similarly, w.r.t. robot P , visible robot coordinates are $R(-3, -4)$, $S(-2, 0)$, and $M(6, -4)$

Next, by adding $(-2, 0)$ with $(3, 4)$, it will get $(1, 4)$. Robot R is already having this coordinate $(1, 4)$ corresponding to robot S , so it will neglect this value. Next, by adding $(6, -4)$ with $(3, 4)$, it will get $(9, 0)$. This is the new coordinate for robot R , so it will store this new coordinate in $\text{Robot}_{\text{Array}}$ and it will update the value of nvr as 3. Now, the values of $\text{Robot}_{\text{Array}}$ for robot R are $S(1, 4)$, $P(3, 4)$, $M(9, 0)$.

Step 3: Next, R will broadcast the updated values of $\text{Robot}_{\text{Array}}$ and nvr along with its flag.

Step 4: Upon receiving the updated values from its neighbor robot(s), the robot R will compare the received nvr values with its own nvr . If all nvr values are equal, then it will execute the next step skipping the next steps. Otherwise, it will repeat the step 2 until all the nvr values become equal. For example, from Fig. 4, after executing step 2, the nvr values of robot R , S , and P will be 3.

Step 5: The robot R will order all the coordinates of $\text{Robot}_{\text{Array}}$ in descending order including itself. If it is the lowest one, then R will set itself as a leader and it will set nf value as number of elements $\text{Robot}_{\text{Array}} - 1$.

Otherwise, it will check its own nf value. If $\text{nf} > 0$, then it will set itself as a semi-leader. Otherwise, it will set itself as a follower. If their x coordinates are same, then their y coordinates compared. As shown in Fig. 4, robot R will store the coordinates in $\text{Robot}_{\text{Array}}$ as $M(9, 0)$, $P(3, 4)$, $S(1, 4)$. Then, robot R will compare its current position coordinate $(0, 0)$ with $S(1, 4)$. Here, $1 > 0$, and finally, the robot R will set itself as a leader and other robots, i.e., S , P , and M are set itself as a follower. Robots S , P , and M store the leader robot coordinate (x_l, y_l) as $(-1, -4)$, $(-3, -4)$, and $(-9, 0)$, respectively.

If the robot R sets itself as a leader, then it will execute the next step and skips the remaining portion of this step.

If the robot R sets itself as a semi-leader or follower, then it will store the leader robot coordinate (l_x, l_y) w.r.t. its own coordinate system for future use and it will execute the *follower* algorithm.

Step 6: The leader robot R will broadcast its current position coordinate to its followers. Once the broadcast to followers is over, the robot R will remain idle until the followers start coming close to it.

After all the follower(s) reach to the coordinate sent by leader, the leader robot will execute the *pathfinder* algorithm.

Algorithm 5 (*follower*)

Step 1: In case robot R detects itself as a semi-leader or follower, then the robot senses the environment. If any new robot(s) is viewed, then it will send its own flag, nvr, and visible robots' coordinate to its new visible robots.

If the semi-leader or follower robot(s) received any instruction, i.e., coordinate from its leader robot, then it will convert this received coordinate w.r.t. its own coordinate system.

Otherwise, follower robot(s) will wait until it receives instruction from its leader.

Step 2: If R is a semi-leader, then it will send this converted coordinate to its follower(s). After sending this coordinate, it will start to move to the newly converted coordinate. Again, R will execute the step 1. Otherwise, it will execute the next step.

Step 3: The follower robot(s) will move toward the newly converted coordinate, i.e., robot P will move to $(-3, -4)$. Again, follower robot(s) will execute the step 1.

4 Conclusion

In this paper, an algorithm is presented to gather all the mobile robots at some point. This algorithm is based on full-compass and asynchronous model. The algorithm guarantees gathering of all the robots within a finite amount of time.

As future works, we would like to extend the algorithm for gathering in such environments.

1. The bounded region may be either convex or concave.
2. The region may or may not consist of obstacles.
3. We can also consider polygonal obstacles.
4. The robots may have limited visibility capabilities.
5. Similar algorithm can be designed based on various timing models such as fully synchronous, semi-synchronous, and various models based on directions and orientations such as half compass, direction only, and axes only.

References

1. Beni, G.: From swarm intelligence to swarm robotics. In: Ahin, E., Spears, W. (eds.) *Robotics Workshop: State-of-the-Art Survey*, no. 3342, pp. 1–9. Springer, Berlin (2005)
2. Cieliebak, M., Flocchini, P., Prencipe, G., Santoro, N.: Solving the robots gathering problem. In: *Automata, Languages and Programming. LNCS*, vol. 2719, pp. 1181–1196 (2003)
3. Cieliebak, M.: Gathering non-oblivious mobile robots. In: *LATIN 2004: Theoretical Informatics. LNCS*, vol. 2976, pp. 577–588 (2004)
4. Agmon, N., Peleg, D.: Fault-tolerant gathering algorithms for autonomous mobile robots. *SIAM J. Comput.* **36**(1), 56–82 (2006)
5. Flocchini, P., Prencipe, G., Santoro, N., Widmayer, P.: Gathering of asynchronous robots with limited visibility. *J. Theo. Comput. Sci.* **337**(1–3), 147–168 (2005)

Design of Wireless Sensor Node to Measure Vibration and Environment Parameter for Structural Health Monitoring Application

Sen Siddheswar, Swarnakar Biplab and Datta Uma

Abstract The paper presents the design of sensor node for wireless sensor network (WSN) dedicated for structural health monitoring (SHM) application. The sensor node, based on Cortex M3 ARM controller, integrates software and hardware for measuring vibration, temperature and humidity data in a compact miniaturized unit. The sensor node is designed for acquiring low frequency vibration with 50 Hz bandwidth. It uses ZigBee protocol based wireless transceiver. A Laptop works as base station to receive information from sensor node. Performance of the sensor node is evaluated by measuring vibration of a 3 phase AC motor and its frequency response is compared with standard vibration data logger Slams StickTM and a single axis mobile industrial accelerometer.

Keywords Wireless sensor network · Structural health monitoring · MEMS · Accelerometer · Temperature and humidity sensor

1 Introduction

Due to prolonged load, fatigue, environment effect and material aging, civil engineering structures such as Bridges, High rise buildings undergo structural deterioration. To avoid the structural collapse, which leads to loss of life as well as monetary loss, the structural health monitoring has become an inevitable preventive measure. SHM systems are capable to estimate the state of structural health and its

S. Siddheswar (✉) · S. Biplab · D. Uma

CSIR—Central Mechanical Engineering Research Institute, Ministry of Science and Technology, Government of India, Durgapur, India
e-mail: s_sen@cmeri.res.in

S. Biplab
e-mail: b_swarnakar@cmeri.res.in

D. Uma
e-mail: umadutta@cmeri.res.in

condition. The most effective way to find out the structural defects is the analysis of vibration data [1]. However Temperature and humidity are also two important environmental parameters which changes structural vibration properties [2]. For the purpose of data communication in SHM system, traditional wired network has been replaced by wireless sensor network (WSN), as the former one experience drawbacks like expensive installation and maintenance [3, 4]. WSN is comparatively new and promising technology which can be implemented in a flexible manner. Reductions of cost, physical size and power consumption are important challenges of WSN system on its path to prosperity. Another challenge related with SHM is small amplitude, low frequency vibration characteristic of structures [5]. With the development of micro-electro-mechanical-systems (MEMS) technology, solution of these challenges has become feasible. ‘Smart Dust’ or Motes are the outcome of the Smart sensing applications development program funded by US Defense Advanced Research Projects Agency under Network Embedded Software Technology program [6]. ‘Rene’ is the first modular construction of mote which was developed in 1999 by CrossBow Technologies [6–8]. The latest available sensor node LOTUS is a product of MEMSIC Inc. It includes two separate boards. The mote board is built around LPC1758 processor and integrates an 802.15.4 compliant radio. To provide flexibility from application point of view, it incorporates a 51 pin connector for attaching application specific sensor board. The results are viewed in Moterview platform. Accelerometer data displayed in the Moterview, represents the instantaneous tilting angle of the mote and does not carry any frequency information [9]. A sensor node named ‘S-Mote’, as described in Yin et al. [4], is developed for Structural Health Monitoring of bridges. Accelerometer, SD1221L having noise density of $5 \mu\text{g}/\sqrt{\text{Hz}}$ at 5 V, is used in ‘S-Mote’. However the sensor draws 8 mA current. Herranen et al. [10] developed a system to acquire and process the acceleration data. The solution includes a sensor board based on MEMS based sensor ADXL326 or ADXL377 and NI data acquisition box USB-6259. The noise density of ADXL326 and ADXL377 are 300 and 250 $\mu\text{g}/\sqrt{\text{Hz}}$ respectively. Acquired data is processed on Labview platform on a PC.

In the present paper, the design of a miniaturized wireless sensor node on a single board, based on LPC1768 Cortex M3 CPU, dedicated for SHM application is described. The node includes accelerometer sensor, temperature and humidity sensor, with a ZigBee firmware loaded wireless transceiver. Vibration, temperature and humidity data are acquired through high speed sampling and sent to the base station through reliable wireless communication. Vibration data is analyzed at the base station using Fast Fourier Transform. Design of a wireless intelligent sensor node for measuring the vibration of structures along with the environment parameters, and processing of signals from sensors for measurement are briefly described in Sect. 2. The efficiency of the developed sensor node for SHM application is verified in Sect. 3. Finally we conclude in Sect. 4.

2 Design of a Wireless Sensor Node and Processing of Signals

The wireless intelligent sensor node is designed by integrating suitable modules/components as shown in Fig. 1. Selection of components is an important part from design perspective. Important features of modules/components to suit design objectives are described in Table 1. The developed sensor node consists of LPC1768 ARM Cortex-M3 based microcontroller [11], Accelerometer ADXL203 [12], humidity and temperature sensor SHT11 [13], a ZigBee protocol based wireless transceiver XBee [14] and a power unit with two AAA battery. System integration is accomplished by the embedded software running on the microcontroller LPC1768, developed on Keil platform. An In-system programming facility of the micro-controller is incorporated through USB interface with proper hardware. The size of the developed sensor node, including battery compartment, is 3 inch × 1 inch × 2 inch.

ADXL203, a dual axis accelerometer, provides ratio metric (560 mV/g for 3 V power supply) analog output signal correspond to acceleration [12]. Outputs of ADXL203 are connected with two 12 bit ADC channels of LPC1768. The voltage reference levels of ADC are set at 3 and 0 V for VREFP and VREFN respectively resulting in a resolution of 1.3 mg (approx). To secure 50 Hz bandwidth, 0.1 uF capacitor is connected at the output of ADXL203 [12]. Temperature and Relative humidity (RH) sensor (SHT11) provide two wire serial interfaces and is connected with LPC1768 through General Purpose I/O pin. The serial clock (SCK) pin and tri state DATA pin of SHT11 are used for synchronization during data communication and bi-directional data transfer respectively. Resolution of the temperature and RH data generated by SHT11 are 14 and 12 bit respectively. Digital readout of

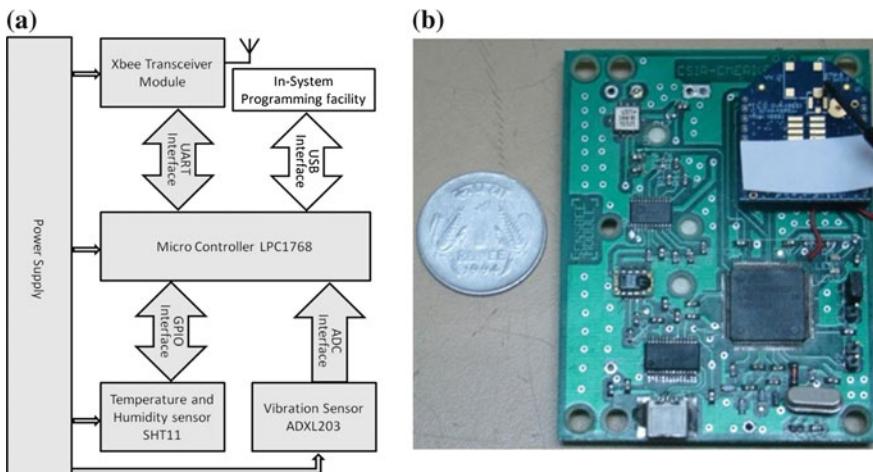


Fig. 1 a Block diagram b developed hardware of wireless sensor node

Table 1 Components/modules used in the design with their important features

Component/module	Feature of interest	Selection criteria
LPC1768 Microcontroller	512 kB flash memory	Eliminate the need of external memory; reduced sensor node size
	64 kB data memory	
	Inbuilt 12-bit ADC	Eliminate the need of external ADC; reduced sensor node size
		Higher bit ADC; increased system resolution
ADXL203	Integrated power management unit	Reduced power consumption
	Low current consumption (450 uA at 3 V)	Reduced power consumption
	Low noise density (190 µg/√Hz rms at 3 Volt)	Enhanced system accuracy
SHT11	Digital output	Interface with microcontroller through general purpose I/O of microcontroller; eliminate the need of external ADC; reduced sensor node size
	Includes sleep mode	Reduced power consumption
XBee	ZigBee protocol	Meets open global standard
	Includes sleep mode	Reduced power consumption

temperature and RH data are converted to physical value using suitable equation and its coefficients as provided in the datasheet [13]. XBEE is an 802.15.4 compliant radio transceiver. Microcontroller interfaces with the XBEE through DOUT and DIN pins for bi-directional data communication. Microcontroller also uses CTS pin of the XBEE for flow control of the data packet send from the microcontroller. The transceiver can be operated as End Device/Router/Coordinator. XBee mounted on the developed wireless sensor node, is configured as ‘End Device’ [14]. The sensor node runs on 3 V power supply, supplied by two AAA size batteries.

The data acquisition from sensors is initiated by data request generated from the base station. On receive the request, the microcontroller of the sensor board acquire temperature and humidity data by issuing suitable command [12], followed by acquiring of vibration data for 4 s, at a sampling rate of 1,250 samples/s. The system requires 20 kB memory for storing dual axis vibration data samples. Vibration data is stored in 250 packets with proper format, each containing 80 byte data appended by 1 byte CRC, and sent to base station through XBee transceiver. Temperature and humidity values are also sent to base station. Flow diagram of sensor data acquisition and successful transmission to the base station is shown in Fig. 2. A laptop is used as the base station. An XBee transceiver, configured as ‘Coordinator’ is connected with the laptop through USB port for communication with sensor nodes. Received data from sensor node is stored and analyzed on

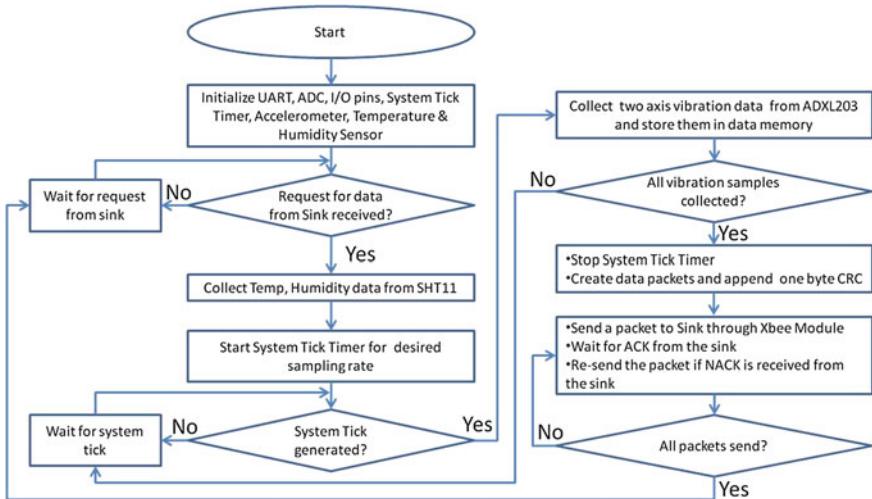


Fig. 2 Flow diagram of data acquisition and transmission from sensor node

MATLAB platform at the base station. Vibration data is plotted in frequency domain after performing FFT on data samples. Temperature and RH values are also displayed.

3 Test Results on 3 Phase AC Motor

Testing of the developed sensor node is performed at the Condition monitoring Laboratory of CSIR-CMERI, by measuring the vibration level of an AC motor excited at 15 Hz (Fig. 3). The speed of the motor is controlled by a Motor controller unit. Complete set of data samples, consisting of vibration, temperature and RH data are received at the base station from the wireless sensor node placed on the shell of the motor structure. Fourier analysis of vibration data is executed using MATLAB FFT function. Amplitudes are further scaled into physical level in the form of ‘g’ by using suitable weight factor (1.3 mg/LSB) and plotted against frequency. This result is validated by the FFT from a vibration data logger (model: Slam StickTM VR-001, make: Mide Technology) [15], acts as reference system and further by using a standard mobile industrial accelerometer having single axis data logging capability (model: VIB 6.142, make: PRÜFTECHNIK Condition Monitoring Germany). Both the reference units are placed on the same structure. Figure 4 shows FFT vibration signal as obtained by the developed system, Slam StickTM and the standard mobile industrial accelerometer (single axis). All responses are showing higher peak response at 30 Hz compared to 15 Hz along with other harmonics. These figures verify the efficacy of the developed system to measure

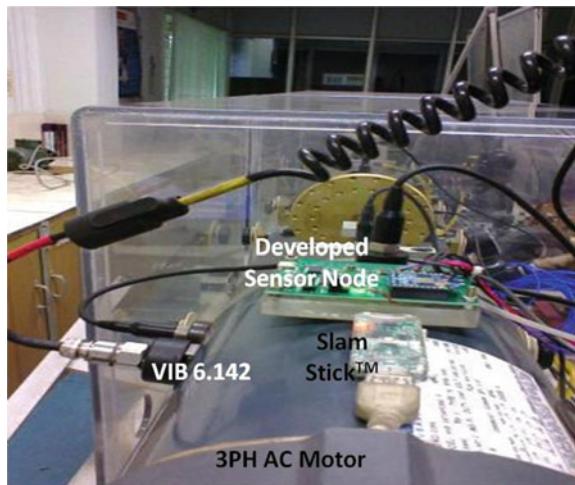


Fig. 3 Test setup for vibration measurement of 3 phase AC motor using developed sensor node, slam stickTM VR-001 and PRÜFTECHNIK VIB 6.142

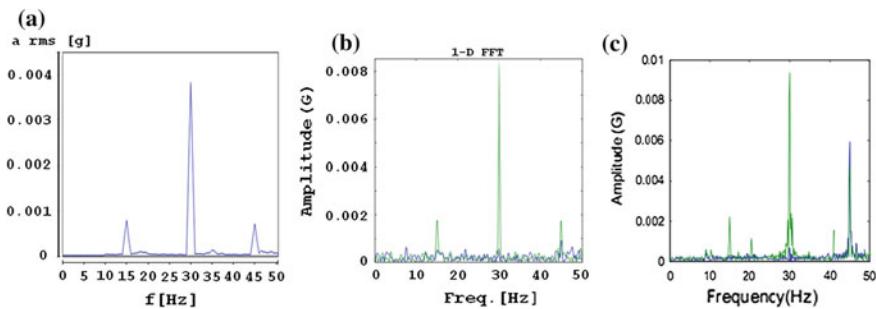


Fig. 4 FFT of vibration data acquired by **a** PRÜFTECHNIK VIB 6.142 **b** slam stickTM VR-001 and **c** developed sensor node

very low level vibration (~ 2 mg). Differences in the value of ‘g’, measured by the systems are observed. It is due to different points of placement as well as different system bandwidth. The temperature and humidity of the environment are also received at the base station and validated by the temperature and humidity meter (Metravi HT-3005).

4 Conclusion

A wireless intelligent sensor node, to detect vibration, temperature and humidity parameters for the SHM of civil engineering structures, is designed. The wireless sensor node will be intelligent in the sense that it preprocesses the complete set of data by suitable algorithm to enhance reliability in communication. Satisfactory results in terms of frequency response of vibration signals are revealed by the experiments and subsequent comparison with the standard data loggers. Temperature and humidity measurements are validated by a standard meter. More tests need to be carried out on actual structure with energy optimal code before practical use in network system.

References

1. Kamrujjaman Serker, N.H.M., Wu, Z.S.: Structural health monitoring using static and dynamic strain data from long-gage distributed FBG sensor. IABSE-JSCE Joint Conference on Advances in Bridge Engineering-II, pp. 519–526, August 8–10 2010
2. Hao, H.: Long-term monitoring of vibration properties of structures with different materials and boundary conditions. In 22nd Australasian Conference on the Mechanics of Structures and Materials, pp. 903–908. Taylor & Francis Group, Sydney Dec 11 2012
3. Notay, J.K., Safdar, G.A.: A wireless sensor network based structural health monitoring system for an airplane. 17th International Conference on Automation and Computing (ICAC), pp. 240–245, Huddersfield, IEEE, 10 Sept 2011
4. Yin, A., Wang, B., Liu, Z., Hu, X.: Design and implementation of the structure health monitoring system for bridge based on wireless sensor network. In: Proceedings of the 6th International Symposium on Neural Networks (ISNN 2009), pp. 915–922, Wuhan, China, 26–29 May 2009
5. Peigneyl, M., Siegert, D.: Piezoelectric energy harvesting from traffic-induced bridge vibrations. Smart Mater. Struct. **22**(9), 1–22 (2013)
6. Spencer, B.F.Jr., Ruiz-Sandoval, M.E., Kurata, N.: Smart Sensing Technology: Opportunities And Challenges. Structural Control and Health Monitoring, Wiley, Hoboken pp. 349–368, 8 Sept 2004
7. Veena Madhuri, V., Umar, S., Veeraveni, P.: A study on smart dust (MOTE) technology. IJCSET **3**(3), 124–128 (2013)
8. Madan, V., Reddy, S.R.N.: Review of wireless sensor mote platforms. VSRD Int. J. Electr Electron. Commun. Eng. **2**(2), 50–55 (2012)
9. www.memsic.com/userfiles/files/Datasheets/WSN/6020-0705-01_A_LOTUS.pdf
10. Herranen, H., Kuusik, A., Saar, T., Reidla, M., Land, R., Martens, O., Majak, J., Acceleration data acquisition and processing system for structural health monitoring. Metrology for Aerospace (MetroAeroSpace), IEEE, pp. 244–248, 2014
11. www.nxp.com/documents/user_manual/UM10360.pdf
12. www.analog.com/static/imported-files/data_sheets/ADXL103_203.pdf
13. www.sensirion.com/.../Sensirion_Humidity_SHT1x_Datasheet_V5.pdf
14. www.digi.com/pdf/ds_xbeezbmodules.pdf
15. www.mide.com/pdfs/slamstick_datasheet.pdf

Economic Load Dispatch Considering Non-smooth Cost Functions Using Predator–Prey Optimization

Sunanda Hazra and Provas Kumar Roy

Abstract In this article, a predator–prey optimization (PPO) is comprehensively developed and successfully applied for solving a single-objective non-linear economic load dispatch (ELD) problem to optimize the generation cost of large-scale power system, considering the non-linearities, such as valve points, and prohibited operation zones (POZs) in the generating units. The effectiveness of the proposed method is examined and validated by carrying out extensive tests on two different test systems, and its results are compared with other those obtained by other techniques available in the recent literature.

Keywords Economic load dispatch • Evolutionary algorithms • Predator–prey optimization

1 Introduction

Due to ever-growing demand for electric energy and since the fuel cost of the power generation is exorbitant, reducing the operation costs of power systems becomes an important topic. The economic load dispatch (ELD) problem is one of the fundamental issues in power systems to obtain optimal benefits and to control the committed generators output such that the total fuel cost is minimized while satisfying the power demand and all other operating constraints [1]. However, practical generators have a variety of non-linearities and discontinuities in their characteristics

S. Hazra (✉)

Electrical Engineering Department, University Institute of Technology,
The University of Burdwan, Burdwan, West Bengal, India
e-mail: sun.hazraee50@gmail.com

P.K. Roy

Electrical Engineering Department, Jalpaiguri Government Engineering College,
Jalpaiguri, West Bengal, India
e-mail: roy_provas@yahoo.com

due to prohibited operating zones and valve-point effects [2], and the cost function is more realistically defined as a segmented piecewise nonlinear function [3].

The ELD problem has been solved via many traditional optimization methods, including gradient-based techniques, Newton methods, linear programming, and quadratic programming. Most of these techniques are not capable of solving efficiently optimization problems with a non-continuous and highly non-linear solution space [4], except dynamic programming [5] in which no restriction is imposed on the shape of cost curves, but this method suffers from dimensionality problem and excessive computational effort.

Recent advances in computation and search for global optimal solution in a finite computation time of complex optimization problems, different artificial intelligence-based methods have been successfully utilized to compute ELD problems. Artificial neural network [6] and intelligent water drop algorithm [7] gained remarkable importance in this domain. Modern heuristic optimization techniques such as simulated annealing (SA) [8], krill herd algorithm [9], genetic algorithm (GA) [10], biogeography-based optimization (BBO) [11], pattern search method (PSM) [12], and gravitational search algorithm (GSA) [13] have been paid much attention by many researchers because of their abilities to find an global solution.

From the literature survey, it is observed that in the past decades, the classical methods had hardly been used to solve ELD problems. Due to several complexities and for advanced computation technique, different artificial intelligence optimization and heuristic methods are incorporated for minimization of the control variable. However, heuristic methods do not always guarantee global best solutions, but are often found to achieve a fast and near-global optimal solution. Recently, Reddy and Vaisakh [14] proposed a hybrid shuffled differential evolution (SDE) algorithm which combined the benefits of shuffled frog leaping algorithm, and DE. Niknam et al. [15] developed new adaptive particle swarm optimization (NAPSO) algorithm because original PSO has difficulties in premature convergence and prevents the convergence to local minima;thus, a new mutation is integrated with adaptive PSO (APSO) by using fuzzy rules. Ciornei and Kyriakides [16] expressed GA-API using ant colony optimization (ACO) and a real-coded genetic algorithm (RCGA) to solve ELD problem. Roy and Bhui [17] implemented quasi-oppositional teaching learning based optimization (QOTLBO) for solving non-linear multi-objective economic emission dispatch (EED). In this paper, an improved predator-prey optimization (PPO) introduced by Li [18] is proposed to solve the non-smooth ELD problem.

The rest of this paper is organized as follows: Section 2 presents the problem formulation. The key points of the proposed PPO technique are described in Sect. 3. In Sect. 4, the proposed method applied to ELD problem is illustrated. Two cases based on large-scale power systems are studied and the simulation results are discussed in Sect. 5. The conclusion is summarized in Sect. 6.

2 Problem Formulations

2.1 Objective Function

The ELD problem is addressed as to minimize the power production cost. Main production cost of thermal generator is fuel cost. Therefore, the objective function of ELD problem may be expressed as follows:

$$F_C = \sum_{j=1}^{NU} F_j(P_j) \quad (1)$$

where F_C is the total fuel cost; $F_j(P_j)$ is the cost of the j th thermal generator, and NU is the number of committed generators to the operating system. The fuel cost $F_j(P_j)$ of the j th thermal generator is computed as a quadratic functions as follows:

$$F_j(P_j) = c_j P_j^2 + b_j P_j + a_j \quad (2)$$

where P_j is the real power output of the j th generator; c_j , b_j , and a_j are the j th generator cost coefficients.

2.1.1 ELD with Valve-point Loading and Prohibited Operating Zones

The sequential valve-opening process for multi-valve steam turbines produces ripple-like effect in the heat rate curve of the generator. This effect is included in ELD problem by superimposing the basic quadratic fuel cost characteristics with a rectified sinusoidal component as follows:

$$F_j(P_j) = c_j P_j^2 + b_j P_j + a_j + |d_j \sin e_j (P_{j,\min} - P_j)| \quad (3)$$

where d_j and e_j are valve-point loading coefficients of generator j . Faults in the generators themselves or in the associated auxiliaries such as boilers and feed pumps may cause instability in certain ranges of generator power output. These ranges are prohibited from operation. The operating zone of the j th unit may be described by

$$P_j = \begin{cases} P_{j,\min} \leq P_j \leq P_{j,1}^{\text{LL}}, \\ P_{j,m-1}^{\text{UL}} \leq P_j \leq P_{j,m}^{\text{LL}}, & m = 2, 3, \dots, Z_m. \\ P_{j,Z_m}^{\text{UL}} \leq P_j \leq P_{j,\max}, \end{cases} \quad (4)$$

where Z_m is the number of prohibited operating zones in the m th generator curve, m is the index of prohibited zone of the j th generator, $P_{j,1}^{\text{LL}}$ is the lower limit of the m th prohibited zone, and $P_{j,m-1}^{\text{UL}}$ is the upper limit of $(m-1)$ th prohibited zone of the j th generator.

2.2 Constraints

2.2.1 Generation Capacity Constraint

For stable operation, power output of each generator is limited by operating limits as follows:

$$P_{j_{\min}} \leq P_j \leq P_{j_{\max}}, \quad j = 1, 2, \dots, \text{NU} \quad (5)$$

2.2.2 Power Balance Constraint

The total power generation must cover the total demand P_{dm} and the real power loss P_{lo} in transmission line. Hence,

$$\sum_{j=1}^{\text{NU}} P_j = P_{dm} + P_{lo} \quad (6)$$

Transmission loss (P_{lo}) of the system is calculated using power flow coefficients by the following formula:

$$P_{lo} = \sum_{j=1}^{\text{NU}} \sum_{i=1}^{\text{NU}} P_j B_{ji} P_i + \sum_{j=1}^{\text{NU}} B_{0j} + B_{00} \quad (7)$$

where B_{00} , B_{0j} , and B_{ji} are loss coefficients which can be assumed to be constant under normal operating condition.

3 Predator–Prey Optimization (PPO)

The PPO technique is based on particle swarm optimization with additional predator chasing property. Predator–prey strategy comes from ecology, and it is inspired by hunting habit of predators to a group of birds (prey), while the preys need to be able to run away from predators. It becomes difficult for preys to stay its preferred location when chased by predators and have to search new location. In this technique, an individual element in prey/predator population represents a solution. Each prey in the population can expand or get killed by predators based on its fitness value, and a predator always tries to kill preys with least fitness in its surroundings, which represents removing (local) of bad solutions and preserving (local) good solutions in the population. The predators are modeled based on the worst solutions as given by (8)

$$(\text{POP})_{\text{predator}} = (\text{POP})_{\text{wrost}} + \delta \left(1 - \frac{t}{t_{\max}} \right) \quad (8)$$

where $(\text{POP})_{\text{predator}}$ is a possible solutions in the populations; $(\text{POP})_{\text{wrost}}$ is the worst solutions in the populations; δ is the hunting rate of the predator; t is the current iteration, and the maximum number of iterations is denoted by t_{\max} . If the predator influences the prey and the interactions between predator and prey provide the solutions to maintain a distance (S) from the predator, then an exponential term will also be included as given by [19].

$$(\text{POP})_{X+1} = \begin{cases} (\text{POP})_X + \delta e^{-|S|}, & S > 0 \\ (\text{POP})_X - \delta e^{-|S|}, & S < 0 \end{cases} \quad (9)$$

where X represents current iterations.

The flow chart of the proposed algorithm is illustrated in Fig. 1.

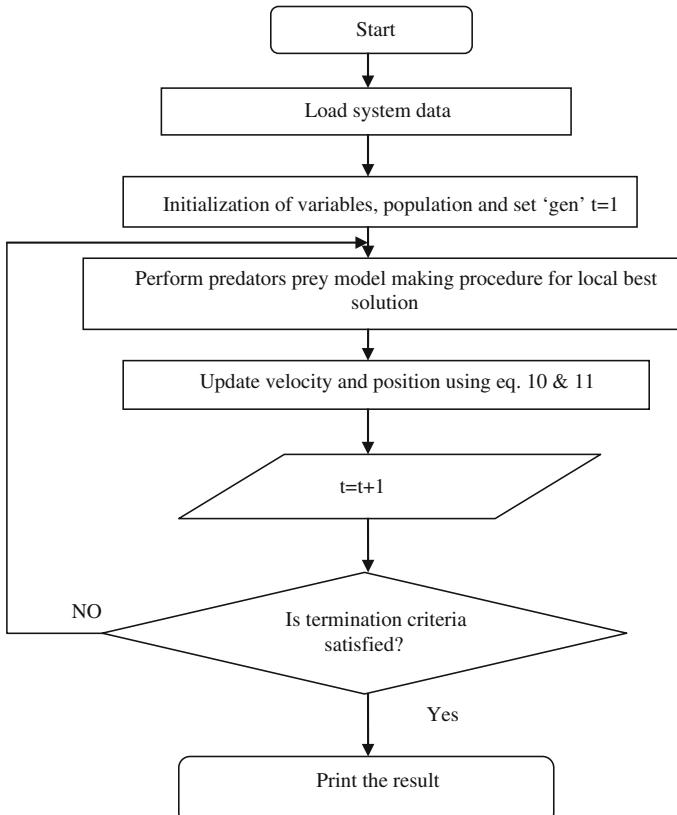


Fig. 1 Flow chart of PPO algorithm

4 PPO Algorithm Applied to Economic Load Dispatch

The following steps must be taken to apply the PPO:

- Step 1: The ELD based on PPO is first begun using input the system data. Specify the generator cost coefficients and valve-point coefficient, choose number of generator units, and specify maximum and minimum capacity constraints of all generators and load demand.
- Step 2: For initialization of the PPO parameter, choose population size P , maximum allowed iterations, and also randomly initialize the prey–predator positions and velocity, individually between upper and lower limit of decision variables. Depending upon the population size, a large feasible solution set is generated.
- Step 3: Compute augmented objective function for all prey population. Assign all prey positions as their local best position based on Eq. (8). Using fuel cost coefficient and generation of each unit, the overall generation cost, i.e., the fitness value is calculated.
- Step 4: Each prey position is evaluated based on fitness value and compute global best position among local best position of prey. Global best position will be the optimum generation of generating unit. The updated velocity and position of a prey and predator at $(t + 1)$ iterations of generating unit are calculated by using the following relation:

$$V_{ij}^{t+1} = wV_{ij}^t + a_1 r_1 \left(P_{\text{best}_{ij}}^t - P_{ij}^t \right) + a_2 r_2 \left(G_{\text{best}_{ij}}^t - P_{ij}^t \right) \quad (10)$$

$$P_{ij}^{t+1} = P_{ij}^t + V_{ij}^{t+1} \quad (11)$$

where $i = 1, 2, \dots, NT$, $j = 1, 2, \dots, NP$. a_1 and a_2 are acceleration constants; r_1 and r_2 are random numbers having value between 0 and 1. $P_{\text{best}_{ij}}^t$ and $G_{\text{best}_{ij}}^t$ are the local best and the global best positions of the i th unit for the j th populations; w is inertia weight.

- Step 5: The influence of the predator on prey is controlled by probability fear, P_f [19]. Then, the solution is found using (9).
- Step 6: This process is repeated for several iterations until the stopping criterion is met.

5 Test Systems and Results

The present work has been implemented in the MATLAB 9 language for the solution of ELD problem. The program was run on CORE™ i3 processor, with 4 GB RAM Personal Computer. For implementing the PPO algorithm, the input

Table 1 Optimal generation and cost obtained by PPO for 40-unit system with loss

Unit	Generation	Unit	Generation	Unit	Generation	Unit	Generation	Unit	Generation
1	113.5367	9	299.9798	17	489.7968	25	523.7182	33	190.0000
2	114.0000	10	279.9786	18	489.3228	26	524.3740	34	200.0000
3	120.0000	11	244.0524	19	511.8249	27	10.1471	35	171.2759
4	180.7326	12	100.9889	20	511.8301	28	10.6337	36	164.8722
5	88.5789	13	482.5948	21	523.4767	29	10.6496	37	110.0000
6	140.0000	14	482.9409	22	528.3479	30	87.8188	38	110.0000
7	300.0000	15	483.8711	23	524.2743	31	190.0000	39	110.0000
8	299.5611	16	483.7963	24	523.1350	32	190.0000	40	549.7983
Cost (\$/h)		136,575,6503							
TL (MW)		969.0084							

parameters used in the simulation study are as follows: population size = 50; number of iterations = 100; probability fear $P_f = 0.3$ and swarm size; inertia (w) and acceleration parameters a_1, a_2 , and $a_3 = 1, 2.5, 1.5$, respectively.

5.1 Test Case 1 (40 Units)

A system with 40 generators with valve-point loading is used here to check the feasibility of the PPO algorithm. The unit characteristics such as cost coefficients along with valve-point loading coefficient and operating limits of generators are given in [3]. Only the transmission losses of 40 units are taken into account to convert it a more realistic practical system. The B-coefficient data of this system are adopted from [16]. The load demand of this system is 10,500 MW. The results obtained after 100 trials, by applying the PPO algorithm and other methods, are summarized in Table 1.

Analyzing the data, it can be observed that the PPO method succeeds in finding a satisfactory solution. This statement is also confirmed by analyzing Table 2, which

Table 2 Statistical results of various algorithms for 40-unit test system with loss

Algorithms	Best fuel cost (\$/h)	Average fuel cost (\$/h)	Worst fuel cost (\$/h)
PPO	136,575.6503	136,576.1246	136,577.3923
KHA [9]	136,670.3701	136,671.2293	136,671.8648
GA API [16]	139,864.96	—	—
SDE [14]	138,157.46	—	—
QOTLBO [17]	137,329.86	—	—
TLBO [17]	137,814.17	—	—

Fig. 2 Cost convergence graph using PPO for 40-unit system

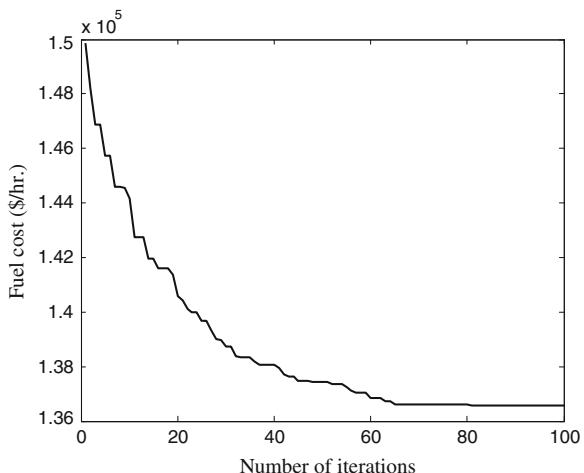


Table 3 Optimal generation and cost of PPO for 80-unit system without loss and with POZ

Unit	Generation	Unit	Generation	Unit	Generation	Unit	Generation	Unit	Generation	Unit	Generation
1	110.8012	18	489.2794	35	194.3974	52	94.0000	69			10.0000
2	110.7998	19	511.2794	36	200.0000	53	214.7598	70			87.7999
3	97.4001	20	511.2794	37	110.0000	54	394.2790	71			190.0000
4	179.7330	21	523.2795	38	110.0000	55	394.2793	72			190.0000
5	87.8000	22	523.2794	39	110.0000	56	394.2801	73			190.0000
6	140.0000	23	523.2794	40	511.2792	57	489.2794	74			164.7994
7	259.5996	24	523.2795	41	110.7998	58	489.2794	75			194.3955
8	284.5998	25	523.2794	42	110.8003	59	511.2792	76			200.0000
9	284.5997	26	523.2794	43	97.3994	60	511.2801	77			110.0000
10	130.0000	27	10.0000	44	179.7324	61	523.2794	78			110.0000
11	94.0000	28	10.0000	45	87.7999	62	523.2794	79			110.0000
12	94.0000	29	10.0000	46	140.0000	63	523.2794	80			511.2790
13	214.7598	30	87.8003	47	259.6002	64	523.2794				242.825.1570
14	394.2793	31	190.0000	48	284.5992	65	523.2794				
15	394.2793	32	190.0000	49	284.6004	66	523.2793				
16	394.2793	33	190.0000	50	130.0000	67					10.0000
17	489.2793	34	164.8001	51	94.0000	68					10.0000

summarizes the best, average, and worst cost obtained by other settled algorithms. From Table 2, the best cost obtained by the proposed PPO method is 136,575.6503 \$/h, which is the best cost found so far. The convergence profiles of fuel cost obtained by the PPO are shown in Fig. 2.

5.2 Test Case 2 (80 Units)

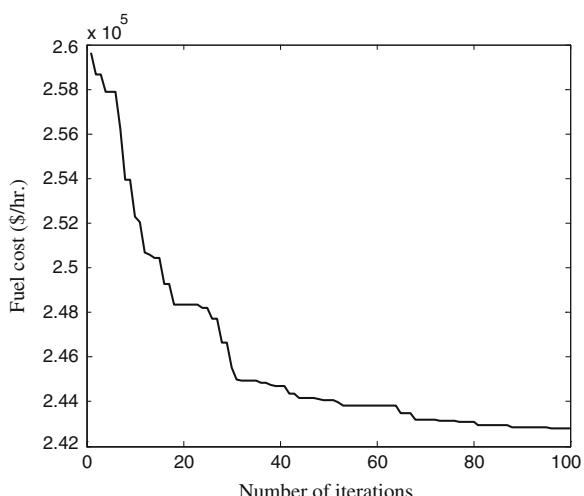
The proposed PPO method is further tested on large-scale systems including lossless 80 units supplying a total load demand of 21,000 MW. The cost coefficients, generation limits, and valve-point coefficients are derived from [3] with modifications to consider prohibited operating zones of units 10–14 and 50–54. The prohibited operating zone data are given in [15]. After performing 100 trials, the best solution obtained by PPO for 80 units is 242,825.1570\$/h, as given in Table 3.

The best, average, and the worst fuel costs obtained by the PPO algorithm are summarized in the Table 4, which shows that the best, average, and the worst fuel

Table 4 Statistical results of various algorithms for 80-unit without loss and with POZ

Algorithms	Best fuel cost (\$/h)	Average fuel cost (\$/h)	Worst fuel cost (\$/h)
PPO	242,825.1570	242,826.3628	242,827.6817
KHA [9]	242,825.2089	242,826.9347	242,828.1350
NPSO [15]	242,844.1172	—	—
FAPSO [15]	244,273.5429	—	—
PSO [15]	249,248.3751	—	—

Fig. 3 Cost convergence graph using PPO for 80-unit system



cost value obtained by proposed method is comparatively less compared to all other methods. A close observation of Table 4 brings out that the difference among the best, average, and worst cost produced by PPO for 100 runs is minimum compared to other methods emphasizing its robustness. The results show that the PPO succeeds in finding the best solution and comparison results explicitly confirm the higher solution quality of the PPO. To illustrate the convergence of the proposed PPO algorithm, over 100 iterations are plotted in Fig. 3.

6 Conclusions

The PPO method successfully accommodates the non-linear characteristics of the cost functions and achieves an optimal solutions. The proposed method has been applied on two test cases with 40 units including losses and lossless 80-unit system. The simulation results so obtained are compared to that of recent approaches reported in the literature. The proposed methodology gives the cheapest generation schedule and outperforms previously reported other methods. Therefore, the proposed PPO method can be a very favorable method for solving the non-convex ELD problem with prohibited zones, especially for the large-scale systems.

References

1. Liu, D., Cai, Y.: Taguchi method for solving the economic dispatch problem with non smooth cost functions. *IEEE Trans. Power Syst.* **20**(4), 2006–2014 (2005)
2. Walters, D.C., Sheble, G.B.: Genetic algorithm solution of economic dispatch with valve point loading. *IEEE Trans. Power Syst.* **8**(3), 1325–1332 (1993)
3. Sinha, N., Chakrabarti, R., Chattopadhyay, P.K.: Evolutionary programming techniques for economic load dispatch. *IEEE Evol. Comput.* **7**(1), 83–94 (2003)
4. Farag, A., Al-Baiyat, S., Cheng, T.C.: Economic load dispatch multi-objective optimization procedures using linear programming techniques. *IEEE Trans. Power Syst.* **10**, 731–738 (1995)
5. Granelli, G.P., Montagna, M., Pasini, G.L., Maranino, P.: Emission constrained dynamic dispatch. *Electr. Power Syst. Res.* **24**(1), 55–64 (1992)
6. Su, C.T., Lin, C.T.: New approach with a Hopfield modeling framework to economic dispatch. *IEEE Trans. Power Syst.* **15**, 541–545 (2000)
7. Rayapudi, S.R.: An intelligent water drop algorithm for solving economic load dispatch problem. *Int. J. Electr. Electron. Eng.* **5**, 43–49 (2011)
8. Erdogan, P., Ozturk, A., Duman, S.: In: International Conference on Electric Power and Energy Conversion System (EPECS'09), Sharjah, pp. 1–4 (2009)
9. Mandal, B., Roy, P.K., Mandal, S.: Economic load dispatch using Krill Herd Algorithm. *Int. J. Electr. Power Energy Syst.* **57**, 1–10 (2014)
10. Yasar, C., Ozyon, S.: Solution to scalarized environmental economic power dispatch problem by using genetic algorithm. *Electr. Power Energy Syst.* **38**(1), 54–62 (2012)
11. Roy, P.K., Ghoshal, S.P., Thakur, S.S.: Biogeography based optimization for economic dispatch. *Electr. Power Compon. Syst.* **38**(2), 168–181 (2009)

12. Sumait, A.L., AL-Othman, A.K., Sykulski, J.K.: Application of pattern search method to power system valve-point economic load dispatch. *Int. J. Electr. Power Energy Syst.* **29**, 720–730 (2007)
13. Roy, P.K., Mandal, B., Bhattacharya, K.: Gravitational search algorithm based optimal reactive power dispatch for voltage stability enhancement. *Electr. Power Compon. Syst.* **40**(9), 956–976 (2012)
14. Reddy, A.S., Vaisakh, K.: Shuffled differential evolution for large scale economic dispatch. *Electr. Power Syst. Res.* **96**, 237–245 (2013)
15. Niknam, T., Mojarrad, H.D., Meymand, H.Z.: Non-smooth economic dispatch computation by fuzzy and self adaptive particle swarm optimization. *Appl. Soft Comput.* **11**(2), 2805–2817 (2011)
16. Ciornei, I., Kyriakides, E.: A GA-API solution for the economic dispatch of generation in power system operation. *IEEE Trans. Power Syst.* **27**(1), 233–242 (2012)
17. Roy, P.K., Bhui, S.: Multi-objective quasi-oppositional teaching learning based optimization for economic emission load dispatch problem. *Int. J. Electr. Power Energy Syst.* **53**, 937–948 (2013)
18. Li, X.: A real-coded predator-prey genetic algorithm for multi-objective optimization. In: *Evolutionary Multi-criterion Optimization (EMO'03)*, vol. 2632, pp. 207–221 (2003)
19. Silva, A., Neves, A., Costa, E.: An empirical comparison of particle swarm and predator prey optimisation. In: *Proceedings of Irish International Conference on Artificial Intelligence and Cognitive Science*, vol. 24(64), pp. 103–110 (2002)

Sparse Denoising in Speckle Noise: A Comparative Study of Dictionaries

Suchismita Maiti, Amish Kumar and Debasish Nandi

Abstract Sparse signal processing appears to be an emerging technology having certain application areas like denoising, deblurring, inpainting, etc. The dictionaries used in sparse signal processing are of much importance as they hold the basic patterns to retrieve the original image. A wide range of complete and overcomplete dictionaries are used for reconstruction of signals in the presence of noise. But no comparative study of these dictionaries is available in the literature till now for indexing their performance. Present work is devoted to carry out such a comparative study which would help in indexing the performance and effectiveness of the dictionaries in sparse signal reconstruction to reduce speckle noise. The results have been compared and analyzed with a set of standard test images.

Keywords Speckle noise reduction · Sparse image processing · Wavelet · Curvelet · KSVD · Overcomplete dictionary

1 Introduction

Image denoising task seems to be more challenging while the problem involves multiplicative noise like speckle noise. Speckle noise is generated due to “coherent radiation” from multiple scatterers which affects a lot of important areas like LASER, SAR, and ultrasound imaging. A number of speckle reduction scheme have already been proposed since last two decades. Among them, some popular methods have been proposed by Kuan [13], Lee [14], Frost [10], Buades [3], Yu [17], Achim [1], etc. Main shortcoming of the proposed speckle reduction techniques is the loss of high-frequency components of the signal, which, in fact,

S. Maiti (✉) · A. Kumar · D. Nandi

Department of Information Technology, National Institute of Technology,
Durgapur 713209, West Bengal, India
e-mail: suchismita2006@gmail.com

introduce the loss of edge information which results in quite good amount of blurring.

During last few years, sparse signal processing has proved its robustness in the field of digital image processing with a wide range of applications like reconstruction, feature extraction, denoising, deblurring, etc. Sparse reconstruction of signal involves a dictionary along with a sparse vector, which is sometimes generated from the noisy signal itself. Dictionary is an important constituent in sparse signal processing as it stores the basis functions, necessary for pattern matching and optimization. The performance of sparse representation-based speckle reduction depends on the nature and type of the dictionary used. Though some literatures on speckle reduction based on sparse signal processing are available, we do not find any suitable comparative study in terms of performance between the dictionaries in the field of sparse signal processing-based speckle reduction so far. In present work, a comparative study on sparse representation-based speckle noise reduction has been performed using different types of dictionaries ranging from off-the-shelf variants to trained dictionaries. The study aims to shed light on the efficiency of different dictionaries in speckle noise reduction using sparse image processing. [9]

This paper is divided into six sections. Sections 2 and 3 contain brief discussion about the statistical model of speckle noise and sparse image processing, respectively. An overview of the dictionaries has been given in Sect. 4. Section 5 contains the details about simulation and the results. The article ends with a conclusion in Sect. 6.

2 Speckle Noise Model

All coherent electromagnetic wave imaging processes face the problem of generation of a typical granular noise, named as speckle noise, due to the presence of multiple scatterers. Speckle noise being signal dependent, non-Gaussian, and spatially independent can be described as one of the most complex noise models. [12] Speckle noise can be mathematically modeled as

$$y(i,j) = x(i,j) \cdot \eta(i,j) + \xi(i,j), \quad (i,j) \in Z^2 \quad (2.1)$$

where speckled image is denoted by $y(i, j)$ and its reconstructed version by $x(i, j)$. Additive noise part [$\xi(i, j)$] can be considered to be negligible as it does not have any significant effect. So, Eq. 2.1 can be rewritten as

$$y(i,j) = x(i,j) \cdot \eta(i,j) \quad (2.2)$$

Hence, Eq. 2.2 can be regarded as the working model for a speckled image.

3 Sparse Representation-Based Denoising

A signal can be modeled as the linear combination of some elementary patterns of a dictionary, termed as atoms, with the help of a sparse vector. More specifically, the product of the dictionary matrix and the sparse vector produces the sparse representation of the signal, which is to be reconstructed, i.e.,

$$D\alpha = I \quad (3.1)$$

where D , α and I correspond to an overcomplete dictionary, a sparse vector and the image, respectively. Structurally, D is a $n \times k$ matrix where $k \gg n$, α is a $1 \times k$ sparse vector. I is a matrix having dimension $n \times b$.

Image denoising problem, when we attempt to solve it using sparse representation modeling, leads to the variant of the optimization problem (P_0^e) as expressed below:

$$(P_0^e) \quad \min_{\alpha} \|\alpha\|_0 \text{ subject to } \|D\alpha - y\|_2^2 \leq \varepsilon \quad (3.2)$$

where the threshold, ε , is closely tied with the noise power, generally ranging from 0.5 to 1.5 and y denotes the noisy image. The solution to this problem, say \hat{x} , is the sparse representation of the desired reconstructed image. Hence, the denoised output would be $D\hat{x}$. The principal advantage of representing a signal in sparse domain lies in obtaining a good accuracy toward an optimum solution of an underdetermined linear equation system. In signal processing area, sparse modeling has been used to the important fields like noise removal, data reconstruction, restoration, etc., as well as in the special fields of image processing like denoising, inpainting, deblurring, etc., though the computation and temporal complexity stills stand as a great hindrance to implement the schemes in real-time systems [9].

4 Overview of Dictionary

The dictionary is an important part of sparse image processing as the reconstructed image is represented as a linear combination of the atoms in the signal space defined by the dictionary. The attributes of “best dictionary” may roughly be listed as the inclusion of fast operators and good candidates, a well-maintained balance between computational time and complexity and the size of the signal space to provide for leading to a sparest representation of the signal. Dictionaries can be of different categories like complete, overcomplete (structurally), and off-the-shelf and learned (according to the building scheme). In present work, the dictionaries are overcomplete, redundant and the variants involve both off-the-shelf and learned ones. Overcomplete dictionaries have more number of atoms than the elements present in the signal which helps in achieving a wide and flexible range of low dimensional signal subspaces. The off-the-shelf dictionaries used here have been generated using Haar, Daubechies, Gabor, DCT, and curvelet transforms [9, 15]. In Haar dictionary,

Kronecker delta function builds the atoms. This transform provides a 50 % compression to the signal. Daubechies transform can be regarded as a generalization of Haar transform. [5] It uses the concept of scaling signal to produce more flexibility in amount of compression. Both Haar and Daubechies transform conserve energy of the signals. The difference between Haar and Daubechies transform lies in the definition of scaling signals, and wavelets DCT can be regarded as an orthonormal transform. The generated coefficients represent frequency contents of the signal. In Gabor transform, a signal is analyzed by taking a windowed portion using a standard window function like Hanning window and applying DFT on it. Gabor atom has been widely used for feature extraction and face recognition in image processing [15, 16]. The efficiency of Gabor transform lies in mainly classification, segmentation, and edge detection tasks. Candes and Donoho [4] proposed “block ridgelet-based transform” for analysis of curvilinear portions of an image, named as curvelet transform. This transform process takes a small portion of an image and applies ridgelet transform on that. Here, we have used KSVD dictionary [2] as an example of learned dictionary which has pros like flexibility to be “conjunction” with any type of pursuit algorithm, simplicity, and its efficiency. KSVD training algorithm incorporates k-means algorithm iteratively along with singular value decomposition approach. The equation below describes the optimization scheme of KSVD:

$$\min_{C,X} \{ ||Y - CX||_F^2 \} \text{ subject to } \forall i, x_i = e_k \text{ for some } k \quad (4.1)$$

where C , X , and e correspond to the original codebook, limited structure of the codebook with k number of entries, and tolerable error, respectively.

Table 1 Result for the dictionaries with 5 sets of test images

Type of dictionary	Barbara			SAR image			Boat		
	Output PSNR	Blur	MSSIM	Output PSNR	Blur	MSSIM	Output PSNR	Blur	MSSIM
DCT	22.45	0.37	0.60	25.89	0.51	0.60	23.2	0.58	0.56
Daubechies	21.58	0.23	0.51	24.05	0.14	0.43	23	0.29	0.53
Haar	21.39	0.26	0.47	22.32	0.15	0.39	23.23	0.29	0.52
Gabor	19.85	0.41	0.44	24.21	0.35	0.53	20.85	0.38	0.45
Curvelet	22.80	0.27	0.59	26.88	0.24	0.65	23.95	0.42	0.57
KSVD	21.72	0.51	0.56	26.04	0.43	0.63	22.7	0.5	0.54
Type of dictionary	Cameraman				Lena				
	Output PSNR	Blur	MSSIM	Output PSNR	Blur	MSSIM	Output PSNR	Blur	MSSIM
DCT	24.32		0.53	0.51	24.18		0.59	0.53	
Daubechies	23.18		0.26	0.53	23.16		0.26	0.50	
Haar	24.19		0.27	0.54	23.03		0.27	0.48	
Gabor	21.15		0.36	0.41	21.1		0.38	0.42	
Curvelet	24.17		0.30	0.44	24.49		0.37	0.52	
KSVD	22.67		0.46	0.49	23.52		0.49	0.52	

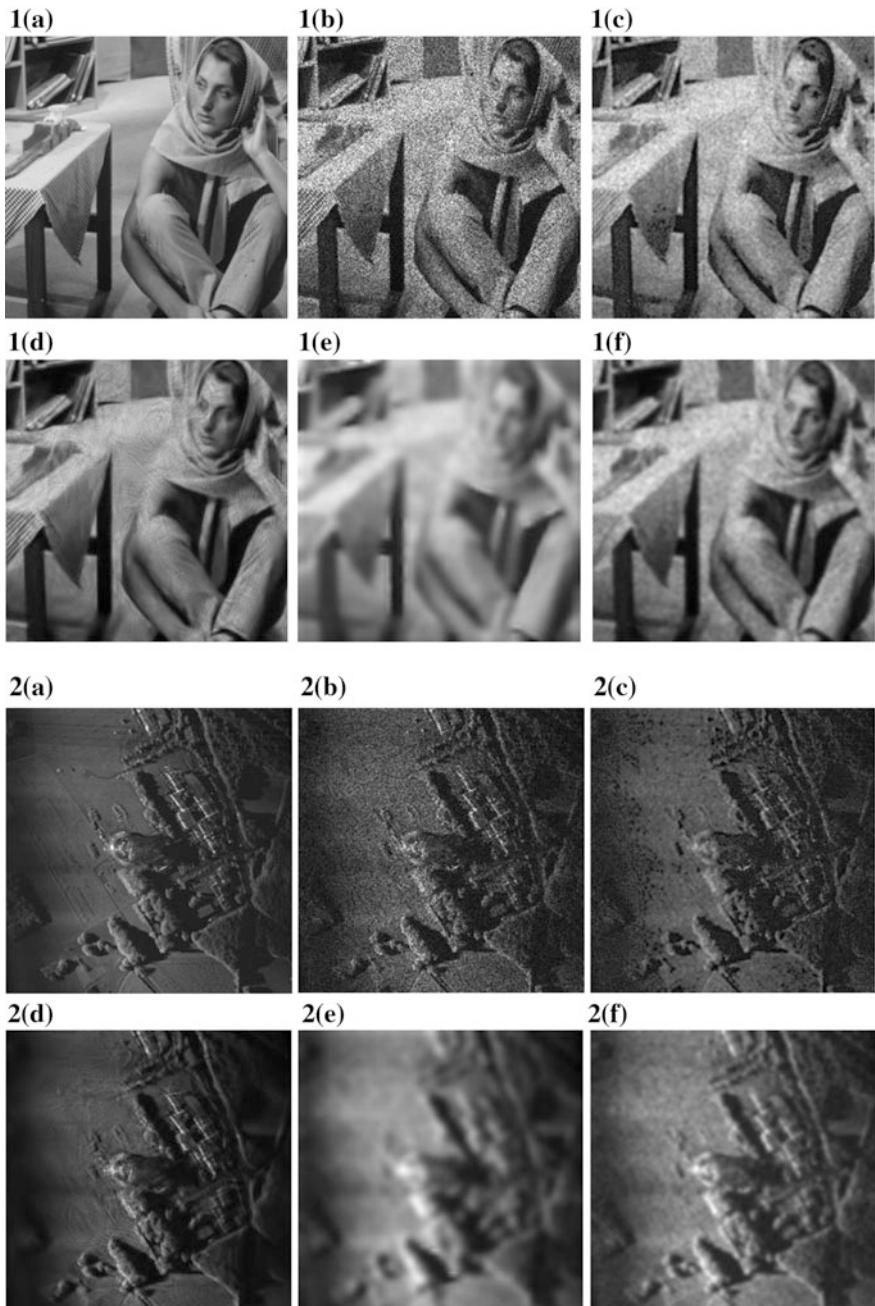


Fig. 1 1–5(a) original images; 1–5(b) noisy images (speckle), 1(c–f) to 5(c–f) Denoised images using sparse image processing with Haar, curvelet, Gabor, and KSVD dictionary, respectively

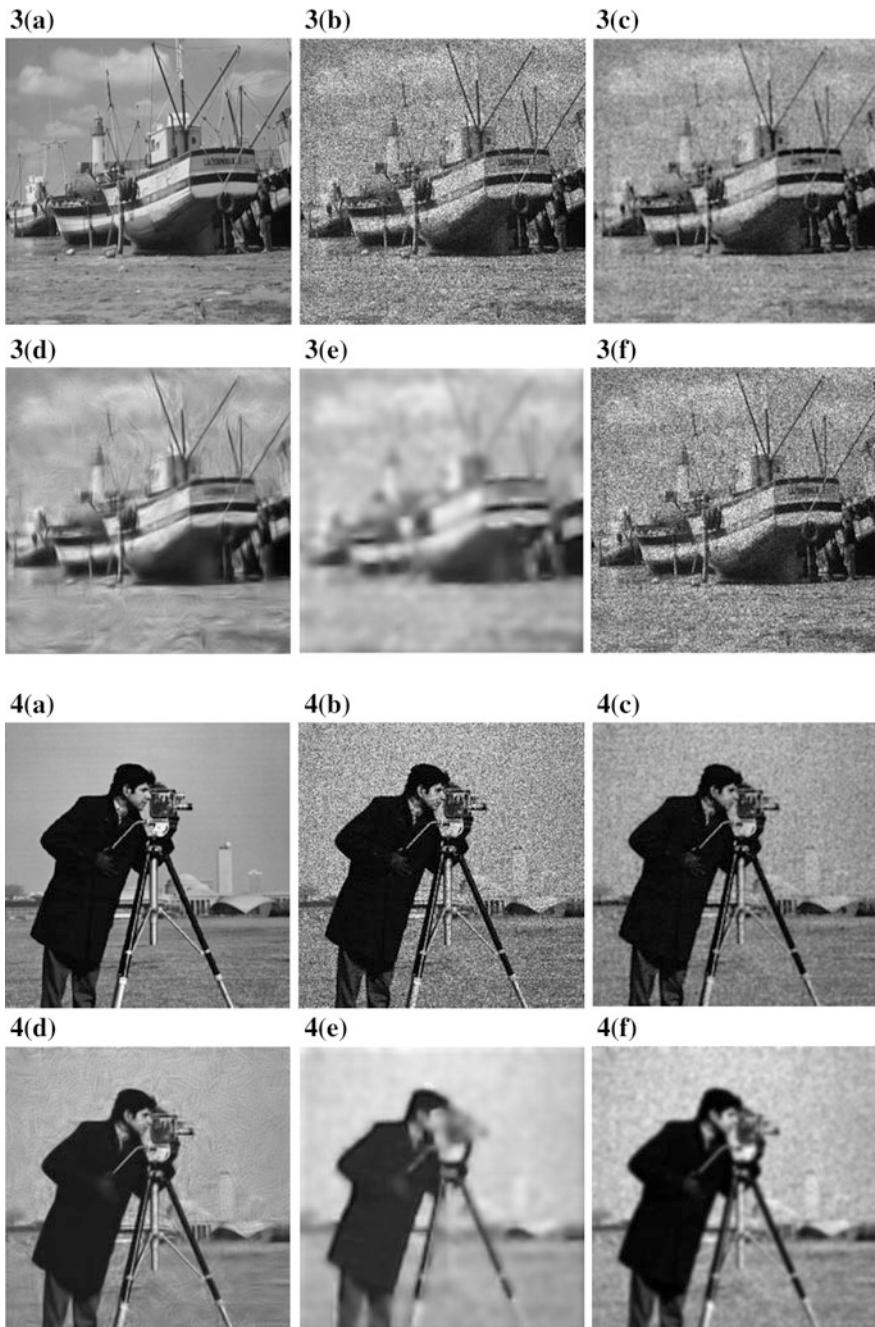


Fig. 1 (continued)



Fig. 1 (continued)

5 Simulation and Result Analysis

All of the experiments have been performed in MATLAB and were simulated on Windows 7 Home Basic, Intel(R) Core(TM)2 Duo CPU@ 3.00 GHz processor with 2 GB RAM, 64-bit OS. General hard thresholding-based sparse denoising schema [9]-based algorithm has been used with six categories of dictionaries, viz. DCT, Daubechies, Haar, Gabor, curvelet, and KSVD. DCT dictionary has been designed with redundancy factor of 4 where the rest of all are having 7. Patch overlapping percentage is 75 %. We have used PSNR, Blur, and SSIM quality metrics for comparison.

Test data set comprises of 5 standard test images like Barbara, Boat, Camerman, Lena, and a SAR image (with good quality). All of the images have been speckled manually with speckle noise having $\sigma = 0.2$ (Table 1; Fig. 1).

The results show different dimensions of despeckling an image with the help of different dictionaries. Curvelet and DCT atoms provide satisfactory value of PSNR though blurring is quite high for DCT. In spite of having a close tie regarding SSIM value, curvelet has an artifact problem compared to other atoms. On the other hand, Daubechies and Haar atoms provide low blurring with a comparatively low PSNR to that of DCT and curvelet. Both Gabor and KSVD produce a high amount of blurring though output PSNR value for KSVD is comparatively fair, especially for SAR image.

6 Conclusion

In present work, the performance of few popular dictionaries has been compared for multiplicative noise reduction using sparse signal processing. The performance has been quantified using few well-defined quality metrics like PSNR, blur metric, and MSSIM. The metrics show that none of the dictionaries could produce satisfactory result in all respect. For example, DCT dictionary has succeeded to produce an excellent amount of noise reduction but suffers from large amount of blurring. For some dictionaries (i.e., Haar and Daubechies dictionaries), the output image is not so blur but they fail to provide an acceptable amount of noise reduction. Further, it is observed that the curvelet dictionary gives birth to some visual artifact which makes the output image unacceptable.

The overall analysis of the results clearly indicates the necessity of designing a more suitable dictionary such that it can balance the amount of noise reduction and the protection of edge information of the image without producing any artifact by its own. In a nutshell, for a proper utilization of sparse signal processing, the dictionary is a key factor which to be designed properly by optimizing denoising efficiency, along with a strong watch on the constraint of blurring issue.

References

1. Achim, A., Tsakalides, P., Bezerianos, A.: SAR image denoising via Bayesian wavelet shrinkage based on heavy-tailed modeling. *IEEE Trans. Geosci. Remote Sens.* **41**, 1773–1784 (2003)
2. Aharon, M., Elad, M., Bruckstein, A.: k-svd: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Process.* **54**(11), 4311–4322 (2006)
3. Buades, A., Bartomeu, C., Morel, J.-M.: A non-local algorithm for image denoising. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005 (CVPR 2005), vol. 2, pp. 60–65. IEEE (2005)
4. Candes, E.J., Donoho, D.L.: Curvelets: a surprisingly effective nonadaptive representation for objects with edges. Department of Statistics, Stanford University, California (2000)
5. Daubechies, I.: The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Inf. Theor.* **36**(5), 961–1005 (1990)
6. Donoho, D.L.: De-noising by soft-thresholding. *IEEE Trans. Inf. Theor.* **41**(3), 613–627 (1995)
7. Candes, E., Donoho, D.L.: Curvelets: a surprisingly effective nonadaptive representation for object with edges. In: Proceeding of Curves and Surfaces IV, pp. 105–121, France (1999)
8. Elad, M., Aharon, M.: Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans. Image Process.* **15**(12), 3736–3745 (2006)
9. Elad, Michael: Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing. Springer, London (2010)
10. Frost, Victor S., Stiles, Josephine Abbott, Sam Shanmugan, K., Holtzman, J.: A model for radar images and its application to adaptive digital filtering of multiplicative noise. *IEEE Trans. Patt. Anal. Mac. Intell.* **2**, 157–166 (1982)
11. Gagnon, L., Alexandre, J.: Speckle filtering of SAR images: a comparative study between complex-wavelet-based and standard filters. In: Optical Science, Engineering and Instrumentation '97, pp. 80–91. International Society for Optics and Photonics (1997)

12. Goodman, J.W.: Some fundamental properties of speckle. *J. Opt. Soc. Am.* **66**(1145–1), 149 (1976)
13. Kuan, D.T., Sawchuk, A.A., Strand, T.C., Chavel, P.: Adaptive noise smoothing filter for images with signal-dependent noise. *IEEE Trans. Patt. Anal. Mach. Intell.* **2**, 165–177 (1985)
14. Lee, J.-S.: Digital image enhancement and noise filtering by use of local statistics. *IEEE Trans. Patt. Anal. Mach. Intell.* **2**, 165–168 (1980)
15. Rubinstein, R., Bruckstein, A.M., Elad, M.: Dictionaries for sparse representation modeling. *Proc. IEEE* **98**(6), 1045–1057 (2010)
16. Lee, T.: Image representation using 2D Gabor wavelets. *IEEE Trans. Patt. Anal. Mach. Intell.* **18**(10), 1–13 (2008)
17. Yu, Y.J., Action, S.T.: Speckle reducing anisotropic diffusion. *IEEE Trans. Image Process.* **11**(11), 1260–1270 (2002)

Compact RS(32, 28) Encoder

Jagannath Samanta, Jaydeb Bhaumik and Soma Barman

Abstract Reed–Solomon codes are commonly used to detect and correct errors in digital data during transmission and storage. In this paper, a new optimization algorithm has been proposed which is very simple and efficient for reducing the complexity of the Galois field constant multipliers in terms of XOR2 gates, and hence, the area overhead of RS(32, 28) encoder decreases. RS(32, 28) encoder has been implemented using four optimized constant field multipliers. Using proposed algorithm, the number of XOR2 gates can be reduced by 34.95 and 50.49 %, respectively, for local and global optimization over non-optimized design without affecting its delay. The number of slices and LUT required for FPGA-based design of RS(32, 28) encoder is also reduced compared to unoptimized design.

Keywords Reed–Solomon code · Galois field multiplier · Compact disk and FPGA

1 Introduction

Error control codes protect digital data against the errors that occur during transmission through a communication channel or during storage. Reed–Solomon (RS) codes are an important class of non-binary block code. RS codes are popularly used for correcting both random and burst errors. RS codes are used to correct errors in many systems including storage devices (CD, DVD), wireless communications,

J. Samanta (✉) · J. Bhaumik

Haldia Institute of Technology, West Bengal University of Technology, Haldia, India
e-mail: jagannath19060@gmail.com

J. Bhaumik

e-mail: bhaumik.jaydeb@gmail.com

S. Barman

Institute of Radio Physics and Electronics, University of Calcutta, Kolkata, India
e-mail: barmanmandal@gmail.com

high-speed modems, and satellite communications [1, 2]. Cross-interleaved RS (32, 28) and RS(28, 24) codes are popularly used in compact disk player [3, 2]. Parr gives two optimization algorithms for Galois Field (GF) multiplication in [4]. An improved method for GF multiplier is proposed which reduces the number of XOR gates to design RS(204, 188) encoder circuit for DVB application [5]. RS encoder based on variable code rates have been presented in [6] for implementing upstream modulation Gbps cable transmission system on different FPGA platforms. In [7], the multipliers in RS encoder that is simple and can achieve high speed of operation have been proposed. A low-complexity and high-efficient RS(255, 223) encoder has been designed in [8]. In [9], some corrections have been done based on the Ref. [8]. In the literature [8, 5], constant multiplier has been implemented using some common sub-expression elimination methods.

Existing optimization techniques are basically based on common sub-expression elimination methods [10]. These methods are very difficult to design RS codec for higher error correction capability. In this paper, we have proposed a new optimization algorithm which is very simple and efficient for reducing the complexity of the GF constant field multipliers in terms of XOR2 gates. Our algorithm can compute the optimized multiplication expressions. We have designed and implemented compact RS(32, 28) encoder for compact disk application. Using proposed algorithm, the number of 2-input XOR (XOR2) gate can be reduced by 34.95 and 50.49 %, respectively, in both local and global optimization compared to unoptimized design. The RS(32, 28) encoder is implemented using 4 optimized constant field multipliers. Proposed RS encoder is simulated and synthesized on FPGA-based Spartan device platform. The number of slices and LUT used is reduced over the non-optimized design. Although CD players are loosing their popularity nowadays, proposed design of double-byte error-correcting encoder still has importance in wireless communication.

The rest of this paper is organized as follows: Proposed construction of constant GF multiplier and its application in RS encoder are discussed in Sect. 2. Section 3 provides the FPGA-based simulation and synthesis results of our proposed design, and then, paper is concluded in Sect. 4.

2 Constant Multiplier in $\text{GF}(2^8)$

The GF adder and GF multiplier are employed to design RS encoder. The GF adder can be implemented easily using XOR gates, but the implementation of multiplier is complicated. Optimized constant GF multiplier is essential for designing a systematic RS encoder. In this section, a new optimization algorithm has been described.

2.1 Proposed New Optimization Algorithm

A new optimization algorithm has been described to reduce the area complexity in terms of XOR2 gates for constant GF multipliers. This algorithm can be used in both local and global optimization.

Proposed Optimization Algorithm It counts the number of XOR2 gates for both local and global optimization. Variables can be defined in the following way:

$$\begin{aligned} m = 8; \quad \text{integer } i, j, k, l; \quad \text{variable } d_{ij}, d_{kl}, d_{ijk}, d_{ijkl}; \\ d_{ij} \leftarrow a_i + a_j; \\ d_{kl} \leftarrow a_k + a_l; \quad d_{ijk} \leftarrow d_{ij} + a_k; \quad d_{ijkl} \leftarrow d_{ij} + d_{kl}; \end{aligned}$$

Step 1: If same combinations of four coefficients are found in two or more equations then replace with a new variable.

Step 2: If same combination of three coefficients are found in two or more equations then replace with a new variable.

```

for  $i \leftarrow 0$  to  $(m - 4)$  do
  for  $j \leftarrow (i + 1)$  to  $(m - 3)$  do
    for  $k \leftarrow (j + 1)$  to  $(m - 2)$  do
      for  $l \leftarrow (k + 1)$  to  $(m - 1)$  do
         $d_{ijkl} \leftarrow a_i + a_j + a_k + a_l;$ 
         $l = l + 1;$ 
      end
       $k = k + 1;$ 
    end
     $j = j + 1;$ 
  end
   $i = i + 1;$ 
end

```

Algorithm 1: Optimization of four variable common sub-expression for step 1

```

for  $i \leftarrow 0$  to  $(m - 3)$  do
  for  $j \leftarrow (i + 1)$  to  $(m - 2)$  do
    for  $k \leftarrow (j + 1)$  to  $(m - 1)$  do
       $d_{ijk} \leftarrow a_i + a_j + a_k;$ 
       $k = k + 1;$ 
    end
     $j = j + 1;$ 
  end
   $i = i + 1;$ 
end

```

Algorithm 2: Optimization of three variable common sub-expression for step 2

Step 3: If same combination of two coefficients are found in two or more equations then replace with a new variable.

```

for  $i \leftarrow 0$  to  $(m - 2)$  do
    for  $j \leftarrow (i + 1)$  to  $(m - 1)$  do
         $d_{ij} \leftarrow a_i + a_j ;$ 
         $j=j+1;$ 
    end
     $i=i+1;$ 
end

```

Algorithm 3: Optimization of two variable common sub-expression for step 3

Step 4: Repeat the step 3, if same combination of two coefficients are found in one equation and replace with a variable which is already used in step 1 and step 2.

Step 5: Rest of the uncommon combinations of coefficients are implemented directly using XOR gates.

Using this algorithm, we can easily determine the total number of XOR2 gates required to design any constant Galois field multiplier circuits. It has been shown that our constant multiplier consumes less area than unoptimized design.

Example Assume **A** is an element in $\text{GF}(2^8)$ and its equivalent binary representation is $(a_7, a_6, \dots, a_1, a_0)$. Let 231 has to multiply with **A** and corresponding output **Y** is $(y_7, y_6, \dots, y_1, y_0)$. After applying the proposed algorithm for local optimization, the bitwise output expressions can be optimized as follows:

$$\begin{aligned} y_0 &= a_6 + a_5 + d_{0123}; & y_1 &= d_{467} + d_{0123}; & y_2 &= d_{467} + a_0; & y_3 &= d_{27} + a_6 + d_3; \\ y_4 &= d_{1467} + d_{25}; & y_5 &= d_{0257} + d_{36}; & y_6 &= d_{1467} + a_3 + a_0; & y_7 &= d_{0257} + a_4 + a_1; \end{aligned}$$

To implement these equations, total 23 XOR2 gates are required, whereas 37 XOR2 gates for unoptimized design.

For global optimization, if the proposed algorithm is used, then the following bitwise expressions are obtained.

$$\begin{aligned} y_0 &= d_{0123} + d_{56}; & y_1 &= d_{0123} + a_4 + d_{67}; & y_2 &= d_{067} + a_4; & y_3 &= d_{67} + d_{23}; \\ y_4 &= d_{67} + d_{1245}; & y_5 &= d_{0235} + d_{67}; & y_6 &= d_{0137} + a_6 + a_4; & y_7 &= d_{0124} + a_7 + a_5; \end{aligned}$$

To implement these equations, 11 XOR2 gates are required and 17 new variables (XOR2 gates) out of 22 are also required (details shown in Table 1).

We have optimized the multiplier complexity in terms of XOR2 gates. Similarly, the other three constant multipliers (116, 216, and 30) are implemented. Using these

Table 1 Complexity comparison of six multipliers for RS(32,28) encoder

Constant multipliers	No. of XOR2 without optimization	No. of XOR2 using proposed local optimization algorithm	No. of XOR2 using proposed global optimization algorithm
116	21	16	6
231	37	23	11
216	13	10	4
30	32	18	8
New variable	0	0	22
Total	103	67	51
Improvement (%) over without opt.	–	34.95	50.49

four multipliers, the RS (32, 28) encoder has been designed, and comparison of complexity is presented in Table 1.

For global optimization, 50.49 % is achieved compared to unoptimized design, whereas 34.95 % improvement is found in local optimization. So the proposed design is very much suitable for VLSI implementation. Design and implementation of RS(32, 28) encoder have been described in the following section.

3 Design and Implementation of RS Encoder

The RS encoder is a circuit which performs polynomial division, and the major problem in designing a small encoder is the large quantity of hardware that is necessary to perform the finite field multiplications [11]. It can be implemented using two different techniques, namely systematic and nonsystematic [12, 13]. In this work, RS encoder has been implemented in FPGA due to its fast and economical way to turn circuit description into hardware. The generator polynomial for double-byte error-correcting code by considering the primitive polynomial $p(x) = x^8 + x^4 + x^3 + x^2 + 1$ in GF(2^8) is as follows:

$$g(x) = x^4 + 30x^3 + 216x^2 + 231x + 116 \quad (1)$$

All the blocks are expressed by using Verilog hardware description language. It has been simulated and synthesized by Xilinx 14.3 ISE simulator, and target device was Spartan3(3s100evq100 with speed grade of 5). Figure 1 shows the architecture of RS(32, 28) encoder. For message block, $m = [1 \ 1 \ 1 \ \dots \ 1 \ 1]$; then, output codeword block is $c = [1 \ 1 \ 1 \ \dots \ 1 \ 1 \ 148 \ 12 \ 24 \ 175]$. Complete synthesis results of optimized RS encoder are given in Table 2. Details of hardware utilization, frequency of operation, and delay are obtained from synthesis report. It has been shown that the area in terms of slices and LUT are reduced by using our

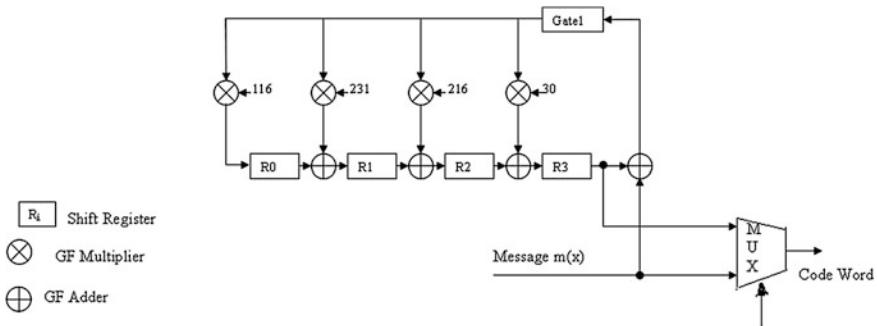


Fig. 1 Architecture of RS(32, 28) encoder

Table 2 Synthesis result of RS(32, 28) encoder using proposed algorithm

	Without optimization	Local optimization	Global optimization
No. of slices (out of 5,472)	79	64	62
No. of slice FFs (out of 10,944)	41	40	40
No. of 4 input LUTs (out of 10,944)	146	124	120
Max. frequency (MHz)	196.33	215.59	187.14
Max. combinational path delay (ns)	6.153	6.153	6.153

proposed algorithm. There is no change in critical path delay of RS encoder. For global optimization, RS encoder area can be reduced up to 21.51 % and the number of LUT used in RS encoder for global optimization has been reduced by 17.81 %.

4 Conclusion

In this paper, a new optimization algorithm has been proposed to design constant $GF(2^8)$ field multiplier. The proposed algorithm is very simple and efficient to reduce the overall complexity of constant GF multipliers. Using this proposed algorithm, the number of XOR2 gates can be reduced by 34.95 and 50.49 % for local and global optimization, respectively. Our design requires less hardware compared to unoptimized design without degrading its speed. In future, we will apply our proposed algorithm to design a higher error correction capability codec such as 16 and 32 bytes error-correcting codes. These optimized constant multipliers are not only used in RS encoder but also can be used in syndrome generation, Chien search and Forney blocks of RS decoder. Overall complexity of the RS codec can be reduced by using these compact $GF(2^8)$ constant multiplier.

References

1. Hanley, S.: Reed-Solomon codes and CD encoding. <http://www.usna.edu/Users/math/wdj/files/documents/reed-sol.htm> (2002)
2. Wicker, S.B., Bhargava, V.K.: Bhargava: Reed-Solomon codes and their applications. IEEE Press, Piscataway (1994)
3. Bhaumik, J., Chowdhury, D.R., Chakrabarti I.: Design and implementation of RS(32, 28) encoder and decoder using cellular automata. In: 15th International Conference on ADCOM-2007, pp. 491–496 (2007)
4. Paar, C.: Optimized arithmetic for Reed-Solomon encoders. In: International Symposium in Information Technology 1997, Ulm, Germany, p. 250 (1997)
5. Wu, X., Shen, X., Zeng, Z.: An improved RS encoding algorithm. In: Proceedings of 2nd IEEE Consumer Electronic Communication and N/W, pp. 1648–1652 (2012)
6. Xiaojun, C., Jun, G., Zhihui, L.: RS encoder design based on FPGA. In: Proceeding of 2nd IEEE, ICACC 2010, vol. 1, pp. 419–421 (2010)
7. Tan, Z., Xie, H., Wu, G., Liao, M.: Design and implementation of Reed-Solomon encoder in CMMB system. In: Proceedings of 6th IEEE International Conference on WiCOM 2010, pp. 1–4 (2010)
8. Dash, R., Lenka, T.R.: VLSI implementation of Reed-Solomon encoder algorithm for communication systems. Int. J. Radioelectronics Commun. Syst. **56**(9), 441–447 (2013)
9. Samanta, J., Bhaumik, J.: Comments on VLSI implementation of Reed-Solomon encoder algorithm for communication systems. Int. J. Radioelectronics Commun. Syst. **57**(7), 331–332 (2014)
10. Jittawutipoka, J., Ngarmnil, J.: Low complexity Reed-Solomon encoder using globally optimized finite field multipliers. In: Proceedings of IEEE Region 10th Conference on Tencon-04, vol. 4, pp. 423–426 (2004)
11. Jinzhou, Z., Xianfeng, L., Zhugang, W., Weiming, X.: The design of a RS encoder. In: Future Computing, Communication, Control and Management. Lecture Notes in Electrical Engineering, vol. 144, pp. 87–91 (2012)
12. Kim, H., Heo, S.W.: Non-systematic RS encoder design for a parity replacer of ATSC-M/H system. IEEE Trans. Consum. Electron. **56**(3), 1270–1274 (2010)
13. Seroussi, G.: A systolic Reed-Solomon encoder. IEEE Trans. Info. Theory **37**(4), 1217–1220 (1991)

Graphene–Silicene Composite can Increase the Efficiency of Cloud Computing

Dhiman Kumar Das and Sukadev Sahoo

Abstract Graphene and silicene are currently attracting the interest of many researchers in the field of material science due to some unique material properties [1]. Recently Reddy et al. [2] have studied the use of using graphene in cloud computing and found graphene is very effective for this purpose, but zero bandgap for graphene is really a problem to use it here. In this paper, we study to improve this property of graphene by preparing a composite of it with silicene and found if graphene is used as an n-type dopant in silicene, then bandgap is induced in graphene. Due to this property, the composite can be used to improve the area of performance, size, speed, and cost in cloud computing services.

Keywords Graphene · Silicene · Composites · Cloud computing

1 Introduction

The structure of graphene consists of two-dimensional sheets of parallel layers of carbon atoms. They are arranged in network of flat regular hexagons. Such layers of carbon skeletons are arranged one above another to form graphite. Each of the carbon atoms of a given layer is joined to its three adjacent carbon atoms in the same layer by using sp^2 hybridization. So to form hexagons, each carbon atoms use three other carbon atoms to form three sigma (σ) bonds, the fourth electron of each carbon atom is used to form pi (π) bond. The C–C bond distance in any layer is

D.K. Das (✉)

Department of MME, National Institute of Technology-Durgapur,
Durgapur, West Bengal, India
e-mail: gournetaidas@rocketmail.com

S. Sahoo

Department of Physics, National Institute of Technology-Durgapur,
Durgapur, West Bengal, India
e-mail: sukadevsahoo@yahoo.com

1.42 Å, and the distance between two adjacent layers is 3.35 Å. The layers are loosely bonded with each other by van der Waals force. Since these structures are loosely linked with each other, they can easily slide over each other which results in their slippery nature and lubricating property. Pi (π) atoms are mobile, hence responsible for conduction of electricity and heat [3].

Silicene is a single-layered 2D structure of silicon analogous to graphene in many respects such as structure, hexagonal arrangements of atoms, and monolayer [4, 5]. But Si–Si bond distance in silicene is 2.28 Å which is much more than C–C bond distances in graphene. In silicene, not only sp^2 but sp^3 hybridization is also observed [5–7]. Silicene also has sigma (σ) as well as pi (π) bonds like graphene, but the pi bond is very weak in nature. This is the cause for its large bond distance compared to graphene [7].

Massless Dirac electrons are observed in graphene, but they are absent in silicene [8–10]. Graphene has no bandgap [2], but silicene possesses a bandgap of 2 meV [6, 11].

Reddy et al. [2] have studied the case of graphene-based transistor in the field of cloud computing. They have found that by using graphene as transistor material, the size of the device will be reduced, electron flow will be much more as electron mobility in graphene is about 1,000 times more than silicon. This property will result in faster operation and increase data transfer rate. Graphene also has the ability to absorb 60 % of the visible light spectrum. They also observed reaction time in graphene is very short about 2.1 picoseconds. Graphene-based sensors allow a frequency of 262 GHz and a data transfer rate over 30 GB/s range of wavelengths when compared with the current commercial networks with a rate of 1.2 GB/s. This kind of performance is based on short lifetime of charge carriers in graphene. Electroluminescence is the result of radioactive recombination observed in graphene confirms its ability to increase internet speed. The power required for off–on of a device made of graphene will be about 70 % less than existing one [2]. Thus, operation cost of devices can be reduced. But a key disadvantage of graphene to be used as transistor is the absence of bandgap. The I_{on}/I_{off} ratios for graphene-based field effect transistors (GFETs) are less than 100, while any successor to the Si MOSFET should have excellent switching capabilities in the range 10^4 – 10^7 [2, 11].

This paper is organized as follows: In Sect. 2, we discuss briefly about cloud computing. In Sect. 3, we show the mechanisms to form graphene–silicene nanocomposite. In Sect. 4, we present our discussions.

2 Cloud Computing

A set of IT services provided to a customer with the ability to scale up or down their service requirements over a network on a leased basis is cloud computing. It is the use of resources available from the internet or computer (hardware or software). Previously network is denoted in the shape of cloud of the sky. A third-party provider who owns the infrastructure usually delivers the cloud computing services.

It advantages to mention but a few include scalability, resilience, flexibility, efficiency, and outsourcing non-core activities. Cloud computing offers an innovative business model for organizations to adopt IT services without upfront investment. Virtualization and abstraction of computing resources bring its wide application [12, 13].

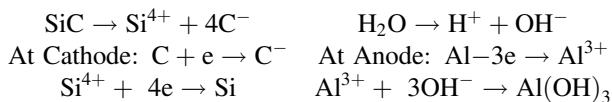
The key characteristics of cloud computing are as follows: User's ability to re-provision technological infrastructure resources is improved with agility. Cloud computing systems typically use representational state transfer (REST)-based APIs. Machine interaction with cloud software in the same way that a traditional user interface (e.g., a computer desktop) facilitates interaction between humans and computers, and their accessibility to software is enabled by application programming interface (API). Cost: cost reduction is a big issue that one cloud claims from providers especially on competitive market. Maintenance of cloud computing applications is easier, because they do not need to be installed on each user's computer and can be accessed from different places. Multitenancy allows for centralization of infrastructure in locations with lower costs (such as real estate, electricity, etc.) by enabling sharing of resources and costs across a large pool of users. Thus, peak load capacity increases (users need not engineer for highest possible load-levels) utilization and efficiency improvements for systems that are often only 10–20 % utilized [14]. Performance is monitored and consistent, and loosely coupled architectures are constructed using web services as the system interface [13, 15]. Multiple users can work on the same data simultaneously rather than waiting for it to be saved and emailed which increases the productivity. Time may be saved as information does not need to be reentered when fields are matched, nor do users need to install application software upgrades to their computer. Reliability improves with the use of multiple redundant sites, which makes well-designed cloud computing suitable for business continuity and disaster recovery. The Virtual Machine (VM) startup time varies by its type, location, operating systems and cloud providers [16–18]. Scalability and elasticity via dynamic (“on demand”) provisioning of resources on a fined-grain is done and self-servicing is provided in urgency by cloud computing [16]. Centralization of data increased security-focused resources, etc., can improve security but loss of control over certain sensitive data, and the lack of security for stored kernels can persist by concerns [13]. There are three types of cloud computing. Private, public, and hybrid cloud computing [12].

3 Mechanism of Graphene–Silicene Nanocomposite Formation

Graphene is the hexagonally arranged 2D nanomaterial of carbon. Graphene is considered as a miraculous material due to its peculiar properties. At room temperature, its electron mobility is $2.5 \times 10^5 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$; Young's modulus of 1 TPa;

intrinsic strength 130 GPa; high thermal conductivity of 4,500 W mK⁻¹ or more [1, 19, 20]. Coefficient of linear expansion for graphene is always negative at least up to 2,500 K [21]. So we can observe graphene is superior than the existing materials. Below, we show a table comparing few properties of graphene with silicene. Atomic number of carbon is 6. sp² hybridization of carbon is shown as 1s²2s²2p_x¹2p_y¹. For silicene, both sp² and sp³ hybridization are observed as 1s²2s²2p⁶3s²3p_x³p_y¹ and 1s²2s²2p⁶3s¹3p_x¹3p_y¹3p_z¹ [3, 22]. So both sigma and pi bonds can be formed between graphene and silicene. Let us consider that the layers are totally ABAB.... stacked one over another and only two layers of the composite is taken into consideration: one is the graphene and other is the silicene (Fig. 1).

Another process may be applied, i.e., electrolysis. Carbon is negatively charged and silicon positively charged. Hence, performing electrolysis process with a carbon cathode and aluminum anode on an acidulated solution of silicon carbide silicon will be deposited on cathode and aluminum hydroxide will be formed at anode (Fig. 2).



The rate of deposition can be controlled by controlling the concentration of solution or intensity of electric current. Only those much minimum parameters are to be provided to deposit single layer of silicene over half the area of the graphene

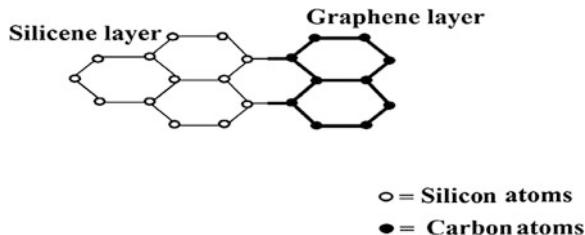


Fig. 1 Graphene–silicene nanocomposite (layers joined in series)

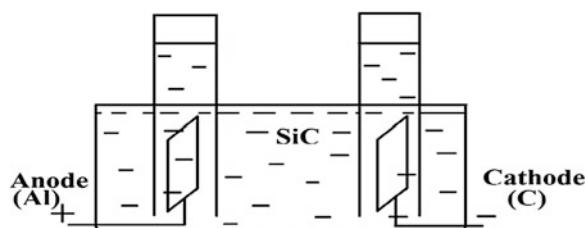


Fig. 2 Electrolysis for preparation of graphene–silicene nanocomposite

cathode at the cost of its own. Thus, only desired nanocomposite can be obtained without formation any bulk substances.

As connected in series, the overall thermal conductivity and electron mobility of the composite will increase but very small in magnitude as both the parameters for silicene is much less than graphene. Considering the equation [22],

$$k_c = k_g + k_s \quad (1)$$

where k_c , $k_g = 4,500 \text{ W/mK}$, and $k_s = 26 \text{ W/mK}$ are the thermal conductivities of the nanocomposite, graphene, and silicene, respectively [23]. Putting these values in Eq. (1), we get $k_c = 4,526 \text{ W/mK}$. Through the processes like heteroatom doping and chemical modification, n-type graphene can be prepared when boron (B) and nitrogen (N) atoms are doped in them [11]. Silicene itself has bandgap of 0.4 eV which can be increased to a maximum value of 0.5 eV by doping alkali metals with it [24]. Now if graphene and silicene are combined to form a graphene–silicene composite, then the composite will have a bandgap. Hence, the main drawback to use graphene as a transistor material, i.e., the absence of bandgap for switch on–off operation could be overcome. Beside these, the mechanical strength of graphene–silicene composite is much stronger than silicene along with high thermal conductivity and bandgap.

4 Discussions

Graphene–silicene nanocomposite has a few modified desirable properties which the parent materials do not alone possess. It will operate faster due to its low reaction time of about 2.1 picoseconds and also have the ability to absorb 60 % of the visible light spectrum. Graphene-based sensors allow a frequency of 262 GHz, and a data transfer rate over 30 GB/s range of wavelengths when compared with the current commercial networks with a rate of 1.2 GB/s. This kind of performance is based on short lifetime (2.1 picoseconds) of charge carriers in graphene. Thus, the speed of internet can be increased. Internet speed and modular frequency each can be increased to a large extent about 1,000 times faster than modern used devices by the stated graphene–silicene nanocomposite. Electroluminescence is the result of radioactive recombination observed in graphene confirms its ability to increase internet speed. This property will result in faster operation and increase data transfer rate. It will improve the cloud computing speed services. The power required for off–on of a device made of graphene will be about 70 % less than existing one [2]. Thus, the transistor on–off problem due to the absence of bandgap can also be rescued by the application of this nanocomposite and also the operation cost of devices can be reduced. But earlier a key disadvantage of graphene to be used as transistor was the absence of bandgap. The $I_{\text{on}}/I_{\text{off}}$ ratios for GFETs are less than 100, while any successor to the Si MOSFET should have excellent switching capabilities in the range 10^4 – 10^7 [2, 11]. But now using silicene with graphene, the

I_{on}/I_{off} ratio is much increased. Hence, the above stated problem becomes a misnomer. So by increasing speed and taking less time to start, the area of performance in cloud computing is increased. The size of transistor can be reduced by using this nanocomposite than previously used silicon. So due to small size, maintenance cost of transistors are reduced, and due to high data transfer rate and quick off-on operation, operational cost is also reduced. High thermal conductivity of graphene–silicene nanocomposite also dissipates heat much quickly than existing material in optical cables which results its performance improvement. Beyond this, the nanocomposite is much more mechanically stable than existing silicon which ensures long life of the device.

Hope our results will be verified experimentally and will bring a drastic change in the field of cloud computing. In the future, we would like to study thermal properties of graphene, try to use n-type semiconductor dopant to open bandgap in it, and try to reinforce graphene with silicene and use them as optical cables with increased numerical aperture, acceptance angle, etc., which will also collect and transmit much more light rays through the cables and thus will increase efficiency of them.

Acknowledgments D.K. Das acknowledges NIT Durgapur for providing me fellowship under “Centre of Excellence (TEQIP II)” scheme.

References

1. Novoselov, K.S., Fal'ko, V.I., Colombo, L., Gellert, P.R., Schwab, M.G., Kim, K.: A roadmap for graphene. *Nature* **490**, 192–200 (2012)
2. Reddy, V.T., Sravan kumar, E., Hasanuddin, M.: Unlocking graphene for faster cloud computing. In: 2nd International Conference on Computer Science and Information Technology, pp. 107–111. Singapore (2012)
3. Maity, S., Ganguly, M.: Elements of Chemistry-1. Publishing Syndicate, Kolkata (2003)
4. Ozcelik, V.O., Cahangirov, S., Ciraci, S.: Stable single-layer honeycomb like structure of silica. *Phys. Rev. Lett.* **112**, 246803–246807 (2014)
5. Voon, L.C.L.Y., Guzman-Verri, G.C.: Is silicene the Next Graphene? *MRS Bull.* **39**, 366–373 (2014)
6. Kaloni, T.P.: Ph.D. Thesis, King Abdullah University of Science and Technology, Thuwal, Makkah Province, Kingdom of Saudi Arabia (2013)
7. Kamal, C., Chakrabarti, A., Banerjee, A., Deb, S.K.: Silicene beyond mono-layers—different stacking configurations and their properties. *J. Phys. Condens. Matter.* **25**(8), 085508–085516 (2013)
8. Zhang, Y., Tan, Y., Stormer, H.L., Kim, P.: Experimental observation of the quantum hall effect and Berry's phase in graphene. *Nature* **438**, 201–204 (2005)
9. Novoselov, K.S., Geim, A.K., Morozov, S.V., Jiang, D., Katsnelson, M.I., Grigorieva, I.V., Dubonos, S.V., Firsov, A.A.: Two-dimensional gas of massless Dirac fermions in graphene. *Nature* **438**, 197–200 (2005)
10. Sahoo, S., Dutta, A.K.: AIP conference. In: Proceedings 1372, pp. 234–237. Melville, New York (2011)
11. Guo, B., Fang, L., Zhang, B., Gong, R.J.: Graphene doping: a review. *Insci. J.* **1**(2), 80 (2011)

12. Kuyoro, S.O., Ibikunle, F., Awodele, O.: Cloud computing security issues and challenges. *Int. J. Comput. Netw.* **3**(5), 247–255 (2011)
13. Wikipedia free encyclopedia for Cloud Computing. http://en.wikipedia.org/wiki/Cloud_computing
14. He, S., Guo, L., Ghanem, M., Guo, Y.: Improving resource utilisation in the cloud environment using multivariate probabilistic models. In: IEEE 5th International Conference on Cloud Computing, pp. 574–581 (2012)
15. Katsaros, G., Kousiouris, G., Gogouvitis, S.V., Kyriazis, D., Menyctas, A., Varvarigou, T.: A self-adaptive hierarchical monitoring mechanism for clouds. *J. Syst. Softw.* **85**(5), 1029–1041 (2012)
16. Mao, M., Humphrey, M.: A performance study on the vm startup time in the cloud. In: 2012 IEEE 5th International Conference on Cloud Computing (CLOUD), pp. 423–430. IEEE (2012)
17. Bruneo, D., Distefano, S., Longo, F., Puliafito, A., Scarpa, M.: Workload-based software rejuvenation in cloud systems. *IEEE Trans. Comput.* **62**(6), 1072–1085 (2013)
18. He, S., Guo, L., Guo, Y., Wu, C., Ghanem, M., Han, R.: Elastic Application Container: A Lightweight Approach for Cloud Resource Provisioning. In: IEEE 26th International Conference on Advanced Information Networking and Applications (AINA), pp. 15–22 (2012)
19. Das, D.K., Sahoo, S.: A provision for using graphene in firewall of aircrafts. In: Proceedings of the Short-term Course on Nanostructured Materials: Theory, Processing and Characterization, p. 26, Durgapur, (2014)
20. Das, D.K., Sahoo, S.: Graphene and it's applications. In: Proceedings of Recent Trends in Condensed Matter Physics, pp. 54–59, Bhubaneshwar (2014)
21. Yoon, D., Son, Y.W., Cheong, H.: Negative thermal expansion coefficient of graphene measured by Raman spectroscopy. *Nano Lett.* **11**(8), 3227–3231 (2011)
22. Dutta, D., Pal, B., Chowdhury, B.: Elements of Physics-1. Publishing Syndicate, Kolkata (2002)
23. Gu, X., Yang, R., arXiv: 1404.2874 [cond-mat.mtrl-sci] (2014)
24. Wikipedia free encyclopedia for Silicene. <http://en.wikipedia.org/wiki/Silicene>

Object-oriented Modeling of IDEA for E-learning Security

Ambalika Ghosh and Sunil Karforma

Abstract E-learning system involves an open participation of student, teacher, and administrator among different regions in the world. Due to the use of Internet as electronic communication media, there are several types of risks and threats such as alteration or deletion of student's marks given by a teacher may hamper the E-learning environment in different ways. To implement privacy and confidentiality of the information, we must use suitable encryption technique. In this paper, we have proposed an object-oriented modeling of international data encryption algorithm (IDEA) for implementation of privacy and confidentiality of information which would be communicated between teacher and student at the time of viewing marks for a subject. For an efficient design, we use UML-based approach.

Keywords E-learning · Encryption · Privacy · Security · Confidentiality · IDEA · Object-oriented model · UML

1 Introduction

The use of information and communication technology (ICT) as a public communicating tool for delivery of services electronically in the public and private sectors has changed the scenario of every operation of the system. This has resulted in emergence of E-learning system [1, 2]. Nowadays, students may not be confined to any conventional school, college, and university campuses only to get their higher education. They may pursue their studies explaining the services of E-Learning [3, 4]. But, most challenging part of E-Learning system is to encrypt the information in public media such as network or Internet [5]. To do so, we have

A. Ghosh (✉) · S. Karforma

Department of Computer Science, Burdwan University, Burdwan, India
e-mail: ambalika_ghosh@yahoo.co.in

S. Karforma

e-mail: dr.sunilkarforma@gmail.com

applied the international data encryption algorithm (IDEA) in object-oriented design. IDEA avoids the use of any lookup tables [6].

E-learning is constructed in a variety of contexts [7], such as online examination, online admit card issue, online payment, and online result that utilize information communication technology (ICT) to promote educational interactions between students, faculties, and administrators [8, 9].

In this paper, we have wrapped IDEA in an object-oriented model for security [10] of information which is needed during view of marks for any subject of a student given by a faculty and published by admin electronically.

In our proposed system, each student is needed to enter his/her roll number, subject code, and subject name to view the marks. Faculty generates the mark slip for the respective subject of the particular student. This passing of information can be possible through admin. This roll number or marks may be altered by hacker causing insecure system. Using IDEA, an encryption–decryption technique is easily possible for protecting the unwanted alternation of roll number or marks from hacker [11].

In Sect. 2, we give a short outline about IDEA. In Sect. 3, we have outlined how object-oriented modeling can be used to implement IDEA for security of information during view marks of a subject. In this paper, we designed and embodied the UML [12–14]-based object-oriented model of E-learning system using IDEA. For an efficient design, we have used use case diagram (Fig. 2), class diagram (Fig. 3), and sequence diagram (Fig. 4) to represent the required class representation for programming purpose in the E-learning system [15, 16] (Fig. 1).

2 IDEA—An Overview

In cryptography [17], the IDEA [6], originally called improved proposed encryption standard (IPES), is a symmetric-key block cipher. The IDEA encryption algorithm provides high-level security not based on keeping the algorithm a secret, but rather upon ignorance of the secret key. It is fully specified and easily understood and also available to everybody. It can be economically implemented in electronic component such as VLSI chip and can be used efficiently. It is also patent protected to prevent fraud and piracy [4].

IDEA operates on 64-bit blocks using a 128-bit key and consists of a series of eight identical transformations (a *round*, see the illustration) and an output transformation (the *half round*). The processes for encryption and decryption are similar. IDEA derives much of its security by interleaving operations from different groups—modular addition and multiplication, and bitwise eXclusive OR (XOR)—which are algebraically “incompatible” in some sense. In more detail, these operators, which all deal with 16-bit quantities, are discussed below [18]:

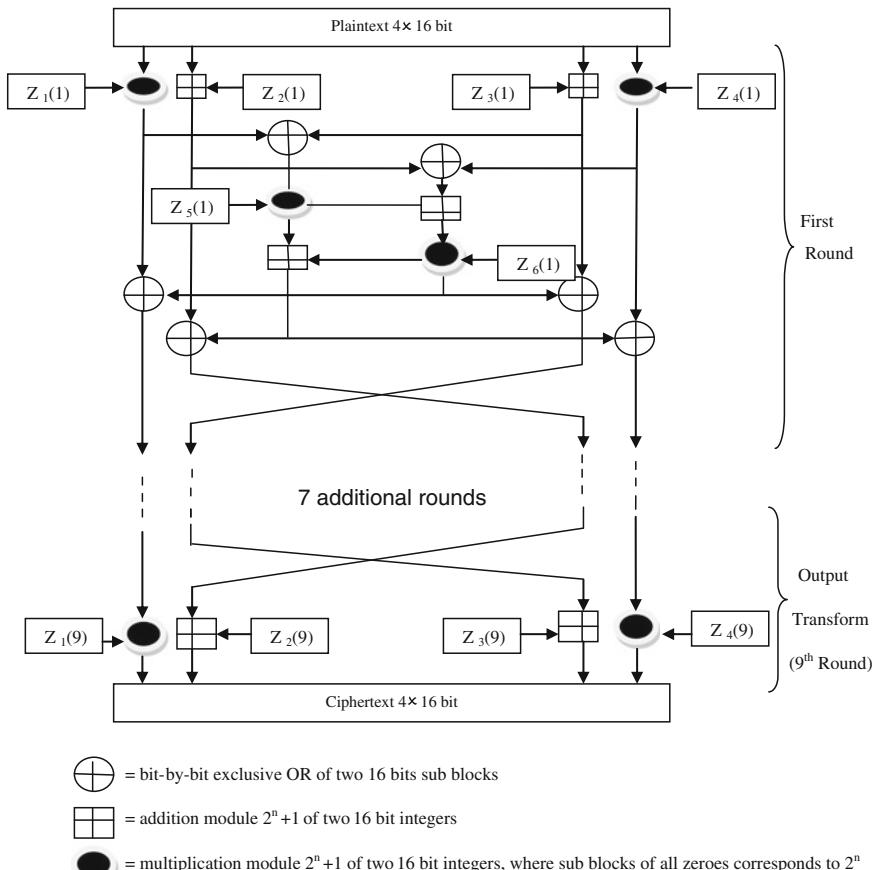


Fig. 1 The IDEA structure

- Bitwise eXclusive OR (denoted with a blue circled plus \oplus).
- Addition modulo 2^{16} (denoted with a green boxed plus \boxplus).
- Multiplication modulo $2^{16} + 1$, where the all-zero word (0×0000) in inputs is interpreted as 2^{16} and 2^{16} in output is interpreted as the all-zero word (0×0000) (denoted by a red circled dot \odot).

After eight rounds comes a final “half round,” this is for output transformation.

3 UML-based Proposed Object-oriented Modeling

To depict our proposed system using UML, we only consider the use case diagram, class diagram, and sequence diagram.

3.1 Use Case Diagram

In the use case model, we can see that there are three types of objects such as student, faculty, and admin who are involved mainly in View Marks use case. In View Marks, let a student enter his/her roll number, subject code, and subject name to view his/her marks in that subject and the faculty provides the marks to him. It is necessary to protect the roll number and marks from any unwanted alteration in the public medium.

This View Marks may be subdivided into three use cases: key scheduling, encryption, and decryption.

In key scheduling, 128-bit key is processed that is used for encrypting the student's roll number and marks given by faculty.

In encryption, the information is encrypted through a repeated process of eight full round and one half round transformations.

In decryption, the encrypted information such as roll number and marks must be decrypted so that user can read correct message.

Use case model of the proposed E-learning system is shown in Fig. 2.

(1) Key scheduling

The 64-bit plaintext block is partitioned into four 16-bit sub-blocks: X_1 , X_2 , X_3 , and X_4 , since all the algebraic operations used in the encryption process operate on 16-bit numbers. Again, six 16-bit key sub-blocks: Z_1 , Z_2 , Z_3 , Z_4 , Z_5 , and Z_6 from the 128-bit key are generated for each of eight rounds. Since a further four 16-bit key sub-blocks are required in the final half round for the subsequent output transformation, a total of $52(=8 * 6 + 4)$ different 16-bit sub-blocks have to be generated from the 128-bit key.

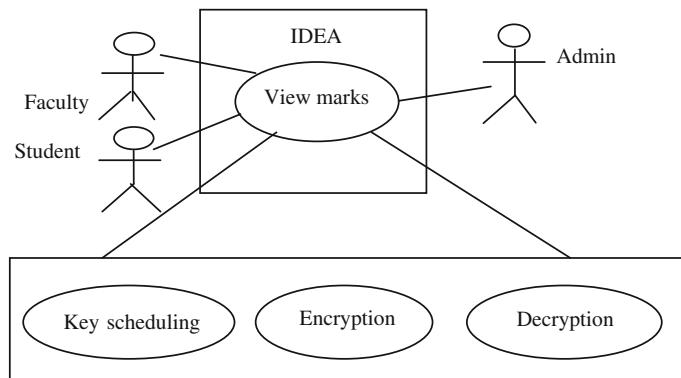


Fig. 2 Use case model of E-learning system

Scenario 1: Mainline sequence

1. IDEA system: Produces 52 16-bit sub-keys from 128-bit key using some steps that are given below:

- First, the 128-bit key is partitioned into eight 16-bit sub-blocks which are then directly used as the first eight key sub-blocks.
- The 128-bit key is then cyclically shifted to the left by 25 positions, after which the resulting 128-bit block is again partitioned into eight 16-bit sub-blocks to be directly used as the next eight key sub-blocks.
- The cyclic shift procedure described above is repeated until all of the required 52 16-bit key sub-blocks have been generated.

(2) **Encryption**

There are fourteen steps of a complete round to encrypt data using IDEA. These fourteen steps will be repeated for eight rounds, and then a half round will be performed [19].

Scenario 1: Mainline sequence

1. Student: Enters his/her roll number, subject code, and subject name for viewing marks to faculty.

2. IDEA System: Accepts information and encrypts it using fourteen steps given below [1]:

- Multiply X_1 and the first sub-key Z_1
- Add X_2 and the second sub-key Z_2
- Add X_3 and the third sub-key Z_3
- Multiply X_4 and the fourth sub-key Z_4
- Bitwise XOR the results of steps 1 and 3
- Bitwise XOR the results of steps 2 and 4
- Multiply the result of step 5 and the fifth sub-key Z_5
- Add the results of steps 6 and 7
- Multiply the result of step 8 and the fifth sub-key Z_6
- Add the results of steps 7 and 9
- Bitwise XOR the results of steps 1 and 9
- Bitwise XOR the results of steps 3 and 9
- Bitwise XOR the results of steps 2 and 10
- Bitwise XOR the results of steps 4 and 10

For every round except the final transformation, a swap occurs and the input to the next round is given as: Concatenation between result of step 11, result of step 13, result of step 12, and result of step 14 which becomes $X_1X_2X_3X_4$, the input for the next round.

After round 8, a ninth half round final transformation occurs:

- Multiply X_1 and the first sub-key Z_1
- Add X_2 and the second sub-key Z_2
- Add X_3 and the third sub-key Z_3
- Multiply X_4 and the fourth sub-key Z_4

The concatenation of the block is the output.

3. IDEA system: Transfer the output encrypted data through the public medium to the admin site.

(3) **Decryption**

The decryption method is same as encryption algorithm except that key scheduling is slightly different. Each of the 52 16-bit key sub-blocks used for decryption is the inverse of the key sub-block used during encryption in respect of the applied algebraic group operation.

Scenario 1: Mainline sequence

1. IDEA system: Accepts the encrypted data from student side and decrypts the data in same way which is used in encryption use case but decryption key is constructed in some reverse fashion.
2. IDEA system: Provides the decrypted original information made by student to admin.

Scenario 2:

1. Admin: Sends the received information to faculty
2. Faculty: Gives the marks according to information

3.2 Class Diagram

The given Fig. 3 demonstrates the organization of class hierarchy showing how a student can view his/her marks of a particular subject given by a faculty with the help of admin using IDEA in object-oriented approach [16].

Class hierarchy of the proposed E-learning system is shown in Fig. 3.

The necessary class diagram for the use of IDEA in E-learning security is given below:

The required four types of classes are: student, faculty, admin, and IDEA system.

Class Student: Some variables are- roll_no, sub_code, sub_name, and methods are- input(), display(), setmsg(), receiveinfo()

Class Faculty: Some variables are- Fid, name, subject and methods are- input(), display(), marks()

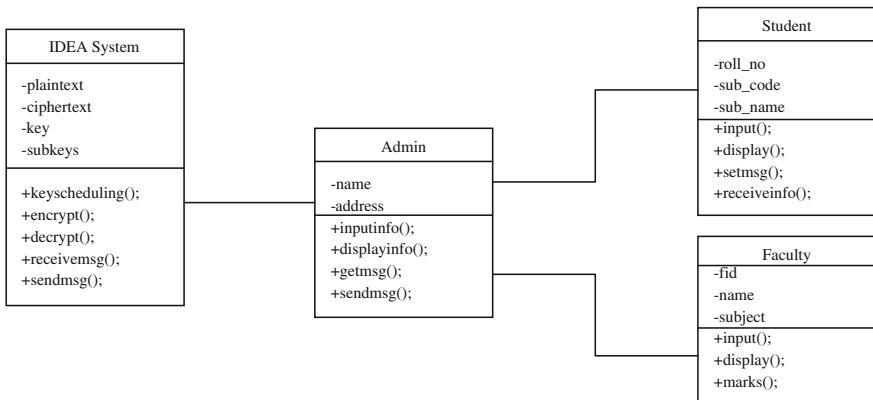


Fig. 3 Class hierarchy diagram of E-learning system

Class Admin: Some variables are- address, name, and methods are- inputinfo(), displayinfo(), getmsg(), sendmsg()

Class IDEA System: Some variables are- Plaintext, Ciphertext, Key, Subkeys and methods are- keyscheduling(), encrypt(), decrypt(), receivemsg(), sendmsg()

3.3 Sequence Diagram

A sequence diagram shows interaction among objects as a two-dimensional chart [11]. Here, we only describe the steps that are needed to encrypt and decrypt any

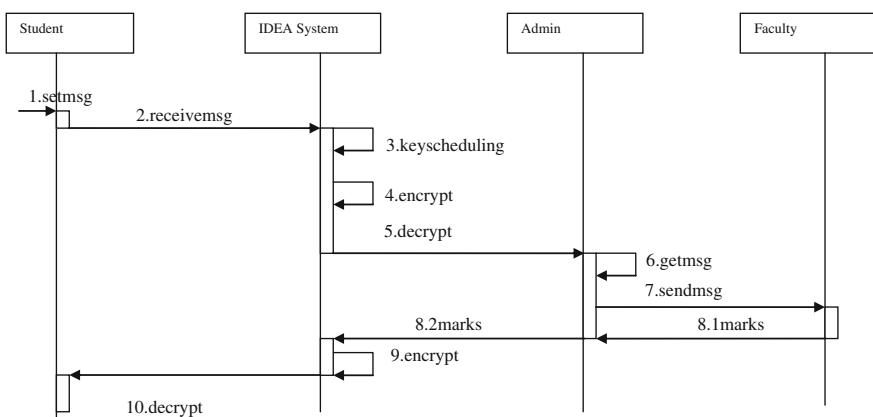


Fig. 4 Sequence diagram of view marks in E-learning system

information providing by student and faculty with the help of administrator during view marks using IDEA algorithm in E-learning system with the help of sequence diagram as in Fig. 4.

4 Conclusion

To ensure the privacy and confidentiality of information, suitable encryption technique is necessary for E-learning security [15, 20]. The proposed system is implementing security of information in E-learning employing IDEA in which encryption and decryption of any content is done very efficiently compared to other algorithms. With implementation of IDEA wrapped in object-oriented model using UML, we can realize a level of safety, reliability and hence trust in the mind of huge number of students. This proposed system will help to reuse the code design [21]. An effortless secure transfer and exchange of data between the faculty and student along with other component of the system become easy. The level of security of the proposed system can be improved further by the efficiency measurement of object-oriented metrics can be considered as the future scope of this work.

References

1. Ghosh, A., Karforma, S., Singh, A.K.: Object oriented modeling of E-learning system. In: Proceedings of ICCS-2010, pp. 103–111. ISBN: 93-80813-01-5
2. Singh, A.K., Mukhopadhyay, S., Ghosh, A., Karforma, S.: Object oriented design of E-library for E-education. In: Proceedings of ICCS-2010, pp. 163–167. ISBN: 93-80813-01-5
3. Weippl, E.R.: Advances in E-Learning, Springer, Berlin
4. Chang, H.S.: International Data Encryption Algorithm, CS-627-1, Fall (2004)
5. Schneier, B.: Applied Cryptography, 2nd edn. Wiley, New York (2008)
6. Roy, A., Banik, S., Karforma, S., Pattanayak, J.: Object oriented modeling of IDEA for E-governance security. In: Proceedings of ICCS-2010, pp. 263–269. ISBN: 93-80813-01-5
7. Eibl, C.J., et al.: Development of E-learning design criteria with secure realization concepts. In: Mittermeir, R.T., Syslo, M.M. (eds.) ISSEP 2008, LNCS 5090, pp. 327–336. Springer, Berlin (2008)
8. Ghosh, A., Karforma, S.: Object oriented modeling of digital certificate based E-learning system. In: Proceedings of RHECSIT-2012, pp. 103–111. ISBN: 978-81-923820-0-5
9. Ghosh, A., Karforma, S.: Object oriented modeling of digital certificate for secure transaction in E-banking. In: Proceedings of NaCCS-2012, pp. 103–111. ISBN: 93-80813-18-X
10. Jamwal, D., Bhat, A.: E-Learning security Concepts. www.google.com/ejel.com
11. Ghosh, A., Karforma, S.: Object oriented modeling of SET for security in E-learning. Int. J. Adv. Res. Comput. Sci. Softw. Eng. 4(8) (2014). ISSN: 2277-128X
12. Mall, R. (ed.): Fundamentals of Software Engineering, 2nd edn. PHI Publications, New Delhi
13. Aggarwal, K.K., Singh, Y.: Software Engineering, revised 2nd edn. New Age International Publishers, New Delhi
14. Hawryszkiewycz, I.T.: Introduction to System Analysis and Design, 2nd edn. PHI Publication, New Delhi

15. Ghosh, A., Karforma, S.: Object oriented modeling of SSL for secure information in E-learning. In: Proceedings of ICCS-2013, pp. 62–66. ISBN-13: 978-9-35-134273-1, ISBN-10:9-35-134273-5
16. Ghosh, A., Karforma, S.: Object oriented modeling of DSA for authentication of student in E-learning. *Int. J. Sci. Res.* **3**(7) (2014)
17. Graff, Jon C.: Cryptography and E-commerce. Wiley, New York (2001)
18. Kahate, Atul: Cryptography and Network Security, 2nd edn. Mc Graw Hill, New York City (2003)
19. C code implementation of IDEA algorithm. Available at <http://www.koders.com/c/fid43C02E2565B7947584D23C36A6C32E198E06C.aspx?s=des>
20. Ghosh, A., Karforma, S.: An UML based design of E-learning system using digital certificate. *Orient. J. Comput. Sci. Technol.* **5**(2), 257–262 (2012)
21. Lafore, R.: Object Oriented Programming in C++, 4th edn. Techmedia, New Delhi

An Improved BAT-Optimized Cluster-Based Routing for Wireless Sensor Networks

Koteswararao Seelam, M. Sailaja and T. Madhu

Abstract The tiny, battery-powered sensor nodes of the wireless sensor networks (WSNs) sense and send reports to a processing center called sink or base station. The sensor nodes require more energy while gathering information for longer durations. This study proposes a protocol heterogeneous in energy which analyzes basic distributed clustering routing protocol low-energy adaptive clustering hierarchy (LEACH) with BAT optimization algorithm to be used for cluster formation and cluster-head (CH) selection. Pipelining is used for packet scheduling. Simulations show that the energy consumption gets reduced significantly.

Keywords BAT optimization · LEACH · Pipelining · Wireless sensor networks

1 Introduction

A wireless sensor network (WSN) is a collection of nodes in a cooperative network [1] with each node having processing capability (one/more microcontrollers, CPUs, or DSP chips), multiple types of memory (program, data, and flash memories), a RF transceiver (a single omnidirectional antenna), and power source (batteries and solar cells) accommodating various sensors and actuators. Nodes communicate wirelessly and self-organize after deployment in an ad hoc fashion. Sensor networks gather information needed by smart environments from buildings, utilities, industrial, home, shipboard, and transportation systems' automation. A sensor network is fast

K. Seelam (✉) · M. Sailaja
UCEK, JNTUK, Kakinada, Andhra Pradesh, India
e-mail: steevan2@gmail.com

M. Sailaja
e-mail: s.maruvada@rediff.com

T. Madhu
SIET, Narsapur, West Godavari, Andhra Pradesh, India
e-mail: tennetimadhu@gmail.com

and easy to install. However, WSNs are more likely to suffer from communication failures at high densities due to contention for shared communication medium. As the routing methods aim to achieve packet routing along optimal paths, link state routing mechanism that suits a multi-channel, multi-interface wireless network is employed here in this study to make the route decision [2]. Though overheads due to link state information exchange are high, route discovery is prompt without additional control overhead like route request packets. Link state routing ensures reliable communication due to its high availability and is also used in military networks. In a link state routing domain, each router originates link state information for link directly connected to the router and floods information to neighbors. Link state routing protocol runs in an autonomous system (intra-domain) and distributes/maintains routing information, facilitating packet forwarding by routers via a best possible path. The distance vector routing protocol locates the shortest paths between nodes in a network through distributed implementation of the Bellman–Ford algorithm [3]. Distance vector routing is used for intra ad hoc network routing by having each network node acting as a router and participating in routing. The routers maintain a routing table listing possible destinations in a network.

Communication between two adjacent hosts/routers is done by a link channel. Data link layer transfers a datagram across individual links by using the flow and error controls. Flow controls are a set of procedures to restrict data that a sender sends before an acknowledgment. A link layer protocol provides flow control to prevent a sending node on one side of a link from overwhelming receiving node on the link's other side. Multiple access protocols such as channel partitioning protocols, random access protocols, and taking-turns protocols regulate node transmission to shared broadcast channels [4]. Like all shared medium networks, medium access control (MAC) protocol is a technique that enables successful network operation by avoiding collisions so that any two interfering nodes do not transmit simultaneously.

Many MAC protocols were developed for wireless voice and data communication networks. Further, while designing the data collection schemes, an effective scheduling mechanism known as pipelining is adopted [5]. Pipelining either increases clock speed (sampling speed) or reduces power consumption at same speed in a digital signal processing (DSP) system. It ensures combining a set of paths to produce longer execution time than selection of the longest path. For example, consider a loop with two paths that takes same number of cycles to execute [6]. One path has a floating-point addition near path beginning, and the other has a floating-point addition near the end. Alternating between paths produces worst-case execution time as there is structural hazard between two floating-point additions. A pipeline sensor network is shown in Fig. 1.

Some protocols use sensor nodes' location information to construct clusters efficiently to achieve energy efficiency. For clustering topology, cluster members send data to cluster-heads (CH), and data aggregation is at CHs, reducing transmission data and saving energy [7]. Additionally, routes are set up in clusters, thereby reducing routing table size stored in individual sensor nodes. Generally, constructing equal-sized clusters is adopted to prolong network life as it prevents



Fig. 1 A pipeline sensor network showing various types of nodes in linear

CHs premature energy exhaustion. Besides, multi-path routing achieves load balancing. Hence, overall network life decreases in sensor network environment as cluster-heads involve an energy-intensive process with energy consumed being greater than non-cluster-head nodes.

Low-energy adaptive clustering hierarchy (LEACH) is a hierarchical routing protocol widely used in WSN. It uses random nodes rotation to be cluster-heads to distribute network energy consumption evenly. LEACH arranges network nodes into small clusters, choosing one as a cluster-head. A node first senses a target and sends information to its cluster-head [8]. The CH aggregates and compresses information received from nodes and forwards it to base station. CH drains more energy compared to other nodes as they have to forward data to base station located far away.

In this study, to overcome LEACH's disadvantages and to improve its efficiency, it is combined with BAT algorithm for cluster formation and CH selection. The rest of this study is organized as follows. Related works in the literature are discussed in Sect. 2. The LEACH and BAT algorithm are described in Sect. 3. Simulation model and numerical results are discussed in Sect. 4. Finally, conclusions are provided in Sect. 5.

2 Related Work

Marriwala and Rathee [9] improved WSN life with regard to increasing live network nodes using a different CH selection approach based on maximum residual energy and minimum distance. An optimal path was chosen between CHs to transmit data to base station.

A multi-hop routing with low-energy adaptive clustering hierarchy (MR-LEACH) protocol was presented by Farooq et al. [10]. To prolong WSN life, MR-LEACH partitioned network into various cluster layers. Each layer's CH collaborates with adjacent layers to transmit sensor's data to base station. Hence, unlike LEACH protocol, MR-LEACH followed multi-hop routing from cluster-heads to base station to conserve energy. Performance evaluation showed that MR-LEACH achieved significantly than LEACH protocol, providing energy-efficient WSN routing.

A simple modification in LEACH protocol clustering algorithm to exploit virtual multiple-input multiple-output (MIMO)-based user cooperation was proposed by Kong [11]. Instead of selecting a single cluster-head at network layer, M cluster-heads suggested in each cluster obtain diversity order of M in long-distance communication.

An analytical method to evaluate energy consumed based on BER curve was presented. Analysis/simulation showed that cooperative LEACH protocol saved huge energy over LEACH protocol with same bit error rate, data rate, delay, and bandwidth requirements. Also, the new approach achieved higher-order diversity with improved spectral efficiency compared to virtual MIMO-based protocols.

Yang et al. [12] used BAT algorithm to solve topology optimization problems. Results showed that the distribution of different topological characteristics like materials could be efficiently achieved. The BAT algorithm was tested through solving nonlinear design benchmarks. Results suggest that BAT algorithm was very efficient in solving nonlinear global optimization problems and topology optimization.

A new cluster-based method that reduced average energy consumption in each round using a new method for data packet transmission to BS was proposed by Khamfroush et al. [13]. Energy consumption was reduced compared to LEACH through modifications including (1) minimization of hop lengths necessary for data packet transmission within clusters, (2) via continuous data aggregation in/out of clusters, and (3) the selection of the nearest CH to BS for transmission of final data packet to BS in each round. Simulation showed the new algorithm to reduce sensor network's average energy consumption compared to LEACH.

Patel et al. [14] emphasized communication routing protocol which impacts network life and scalability. LEACH is a sensor network routing protocol for energy conservation. Simulation revealed node density performance and cluster-head percentage of LEACH. Scalability and cluster variation improved network life by analyzing with regard to live node, simulation time, energy, and data.

Zhang et al. [15] presented weighted spanning tree algorithm based on LEACH (WST-LEACH). First, CH selection was not completely random; second, it established a weighted spanning tree through all CH, and the weight value calculation of weighted spanning tree contained factors such as CH remaining energy, surrounding nodes distribution, and distance to other CH. The data were sent to base station with this tree after integration. It optimized data transmission path and reduced energy consumption. Simulation results proved that WST-LEACH reduced energy consumption and had higher efficiency and extended network life.

A low-energy consumption chain-based routing protocol LEACH-CC based on LEACH protocol research was proposed by Li and Zhang [16]. The new protocol had each node sending information about current location and energy level to base station. The latter ran simulated annealing algorithm to determine clusters for that round. Then, chain routing between clusters reduced nodes that communicate with base station. Finally, the results showed that LEACH-CC performed better than LEACH. It not only extended network life but also improved energy efficiency.

Ran et al. [17] improved LEACH protocol using fuzzy logic (LEACH-FL), which considers battery level, distance, and node density. LEACH is a famous clustering mechanism which elects a CH based on probability. Nodes energy is an important consideration as WSN's life is limited by nodes energy. The proposed method made a better selection by comparing simulations using MATLAB.

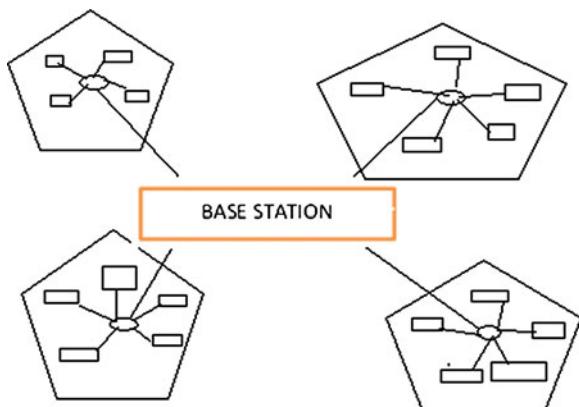
A new meta-heuristic method, BAT algorithm, based on bats' use of echolocation was proposed by Yang [18]. The author planned to combine existing algorithms' advantages into the BAT algorithm. After detailed formulation and explanation of implementation, this algorithm was compared to other algorithms, including genetic algorithms and particle swarm optimization (PSO). Simulations revealed the new algorithm to be superior to other algorithms.

Saha et al. [19] proposed opposition-based BAT algorithm. The simulation results demonstrate the efficiency of BAT algorithm when compared to other evolutionary algorithms as differential evolution (DE), real-coded genetic algorithm (RGA), and PSO in terms of accuracy and convergence speed.

3 Methodology

Using adaptive clusters and rotating cluster-heads, LEACH outperforms classical clustering algorithms, allowing system energy requirements to be distributed among all sensors [20]. Additionally, LEACH performs local computation in each cluster to reduce data to be transmitted to base station, achieving large reduction in energy dissipation, as computation is cheaper than communication. Nodes organize themselves into local clusters, with one node acting as cluster-head in LEACH as shown in Fig. 2. All non-cluster-head nodes transmit data to cluster-head, while cluster-heads receive data from all cluster members, perform signal processing functions on data (data aggregation), and forward it to remote base station. Hence, a cluster-head node is more energy intensive than a non-cluster-head node. LEACH operation is broken up into rounds, with each round beginning with a setup phase, when clusters are organized, followed by a steady-state phase, where data transfers to base station occur. To minimize the overhead, steady-state phase is long compared to setup phase. Initially, when clusters are created, a node decides on becoming a cluster-head for the present round.

Fig. 2 LEACH protocol architecture



LEACH cluster-head selection is stochastic. To select the cluster-heads, a node n determines a random number between 0 and 1. When number is below a threshold $T(n)$, the node becomes a cluster-head for the current round. Threshold setting is as follows [21]:

$$T(n) = \frac{P}{1 - P[r \bmod \left(\frac{1}{P}\right)]} \quad \forall n \in G \quad (1)$$

where P is the cluster-head probability, r the number of current rounds, and G a set of nodes that have not been cluster-heads in the last $1/P$ rounds. This algorithm ensures that all nodes become cluster-heads once within $1/P$ rounds.

LEACH achieves over a factor of 7 reduction in energy dissipation compared to direct communication and a factor of 4–8 compared to minimum transmission energy routing protocol [22]. Nodes die randomly, and dynamic clustering increases system life. LEACH is completely distributed, needing no global network knowledge. However, LEACH uses single-hop routing where each node transmits directly to cluster-head and sink. Hence, it does not suit networks deployed in large regions. In LEACH, an optimal percentage p_{opt} of nodes become CH in each round. Some CH may be very close to each other or located in the edge of the WSN due to which energy efficiency cannot be maximized. In this paper, a method has been proposed for the improvement of LEACH by using BAT optimization method for cluster formation and CH selection based on distances and number of CHs. BAT algorithm uses a frequency-tuning technique which increases the diversity of the solutions in the population and also uses the automatic zooming to balance exploration and exploitation during the search process. Comparison study of BAT algorithm shows certain advantages over PSO, genetic algorithm, and other algorithms [23].

Based on bats' behaviors, Nakamura et al. [24] developed a new meta-heuristic optimization technique called BAT algorithm. This technique was developed to behave as a band of bats tracking prey/foods using echolocation. Echolocation is a form of sonar: Bats, mainly microbats, emit a loud/short pulse of sound and wait for it to hit an object, and after a fraction of time, the echo returns to its ears.

Based on the above BAT echolocation description and characteristics, Yang and He [25] developed BAT algorithm with the three rules:

1. Bats use echolocation to sense distance, and they 'know' the difference between food/prey and background barriers in some way.
2. Bats fly randomly with velocity v_i at position x_i with frequency f_{\min} varying wavelength λ and loudness A_0 when searching for the prey. They adjust wavelength (or frequency) of emitted pulses automatically.
3. They adjust rate of pulse emission $r \in [0, 1]$ depending on the proximity of target; though loudness varies in many ways, it is assumed that loudness varies from a large (positive) A_0 to a minimum constant value A_{\min} .

BAT algorithm (BA) is as follows [26]:

3.1 Algorithm 1: BAT Algorithm

```

Objective function  $f(x)$ ,  $x = (x_1, x_2 \dots x_d)^T$ 
Initialize the BAT population  $x_i$  and  $v_i$  for  $i = 1 \dots n$ 
Define pulse frequency  $Q_i \in [Q_{\min}, Q_{\max}]$ 
Initialize pulse rates  $r_i$  and the loudness  $A_i$ 
While ( $t < T_{\max}$ ) // number of iterations
  Generate new solutions by adjusting frequency, and
  updating velocities and locations/solutions
  if ( $\text{rand}(0, 1) > r_i$ )
    Select a solution among the best solutions
    Generate a local solution around the best solution
    end if
    Generate a new solution by flying randomly
    if ( $\text{rand}(0, 1) < A_i \& f(xi) < f(x)$ )
      Accept the new solutions
      Increase  $r_i$  and reduce  $A_i$ 
    end if
    Rank the bats and find the current best
  end while
Post process results and visualization

```

BA mimics bats when catching prey. BA enforcement is complicated compared to other meta-heuristic algorithms [27] as each agent (BAT) is provided a set of interacting parameters such as velocity, position, loudness, pulse rate, and frequencies which affect solution quality and time needed for such a solution. In the proposed optimization, the fitness is based on the total transmission distance and the number of CHs. The fitness is computed as follows:

$$f = \alpha * (D_{\text{total}} - d_i) + (1 - \alpha) * (N_{\text{total}} - N_{\text{ch}}) \quad (2)$$

where α is the predefined weight, D_{total} is the distance of all nodes to sink, d_i is the sum of distance of nodes to CH, N_{total} is the number of nodes in the network, and N_{ch} is the number of CHs. The fitness of the node is higher if distance is shorter or if CHs are lesser. During initialization, the algorithm randomly selects nodes to be CH in the network. Based on the fitness function, the algorithm finds appropriate number of CHs and its location. The steps in the proposed algorithm are as follows:

3.2 Algorithm 2: Proposed BAT Algorithm

```

Initialize the parameters (population size, frequencies, and the maximum iterative count)
Create a swarm with P bats
Calculate the cluster centers and distances

```

Calculate the fitness value of each BAT
 Update the velocity matrix for each BAT
 Update the location vector for each BAT
 Calculate the (global) best and (local) personal best location for all the bats and each BAT
 If terminating condition is not met, go to step 3.

4 Experimental Results

Simulations are conducted for throughput and media access delay, retransmission attempts, and data dropping for different methods. Table 1 tabulates the parameters of the BAT algorithm. Figure 3 shows the throughputs achieved in experiment.

From Fig. 3, it can be seen that proposed cluster-based pipelining with optimization increases throughput of the system when compared with LEACH.

Table 1 Simulation

parameters of BAT algorithm

BAT population	20
Pulse frequency interval	[0.0, 2.0]
Pulse emission	[0, 1]
Loudness	0.5
Maximum number of generations	1,000

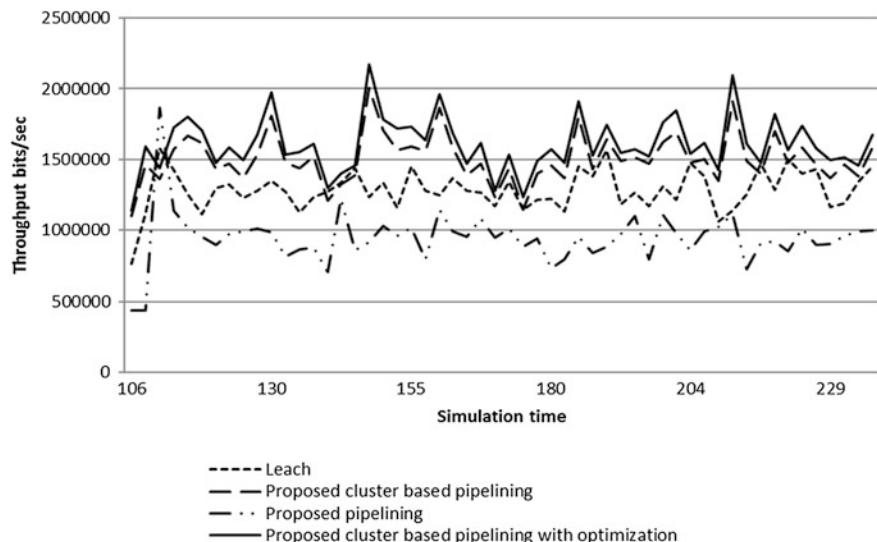


Fig. 3 Throughput versus simulation time

The proposed method improves throughput by an average of 20.19 % than LEACH, by 6.38 % than cluster-based pipelining, and by 41.29 % than pipelining. Figure 4 shows the comparison of delays occurred when varying the number of data channels. From Fig. 4, it can be seen that proposed cluster-based pipelining with optimization method decreases the delay of the system when compared with LEACH. The proposed method's delays are less by an average of 17.61 % than LEACH, by 4.47 % than cluster-based pipelining, and by 24.72 % than pipelining. Figures 5 and 6 depict the retransmission attempts and data dropped.

From Fig. 5, it can be seen that proposed cluster-based pipelining with optimization method decreases the retransmission attempts in the system when compared with LEACH. The proposed method's retransmission attempts are less by an average of 10.08 % than LEACH, by 4.58 % than cluster-based pipelining, and by 27.01 % than pipelining.

From Fig. 6, it can be seen that the proposed cluster-based pipelining with optimization method decreases the data dropped in the system when compared with LEACH. The proposed method's data dropped are less by an average of 20.97 % than LEACH, by 6.37 % than cluster-based pipelining, and by 31.27 % than pipelining.

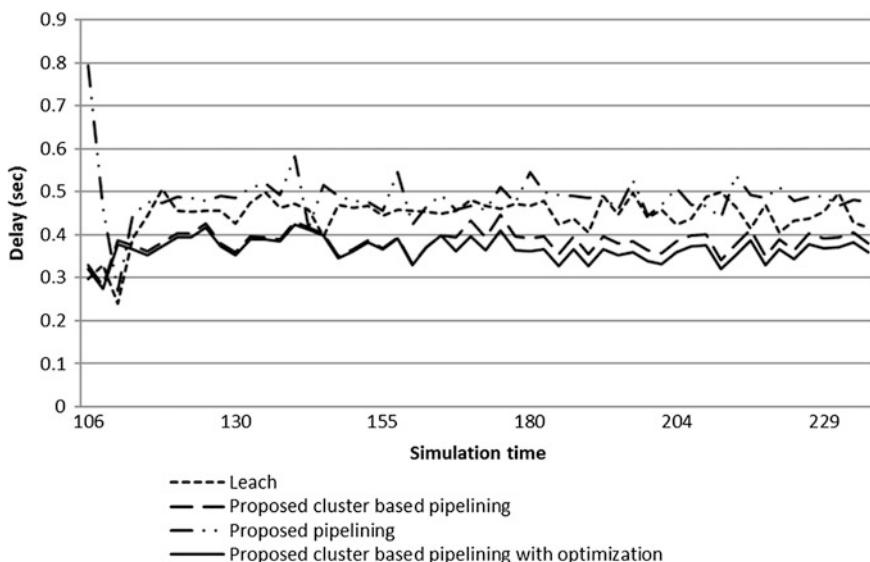


Fig. 4 Delay versus simulation time

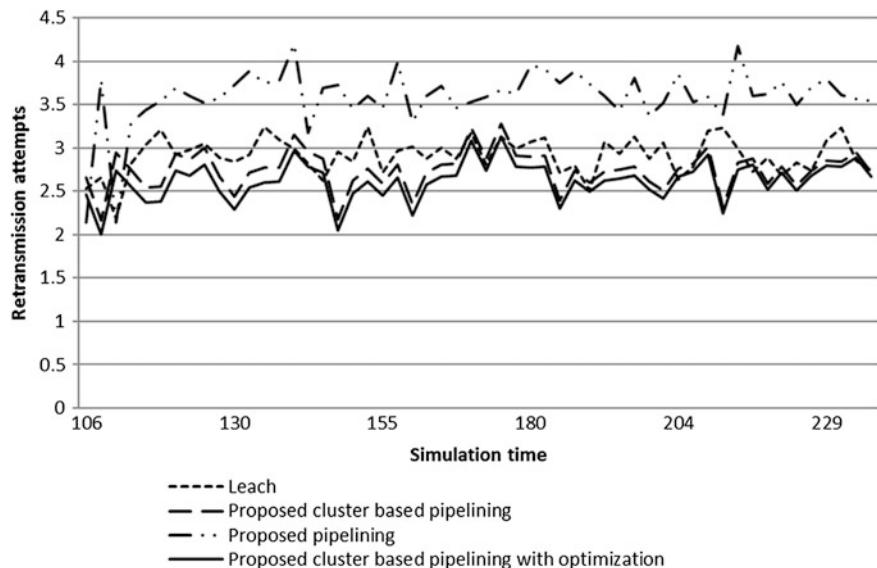


Fig. 5 Retransmission attempts versus simulation time

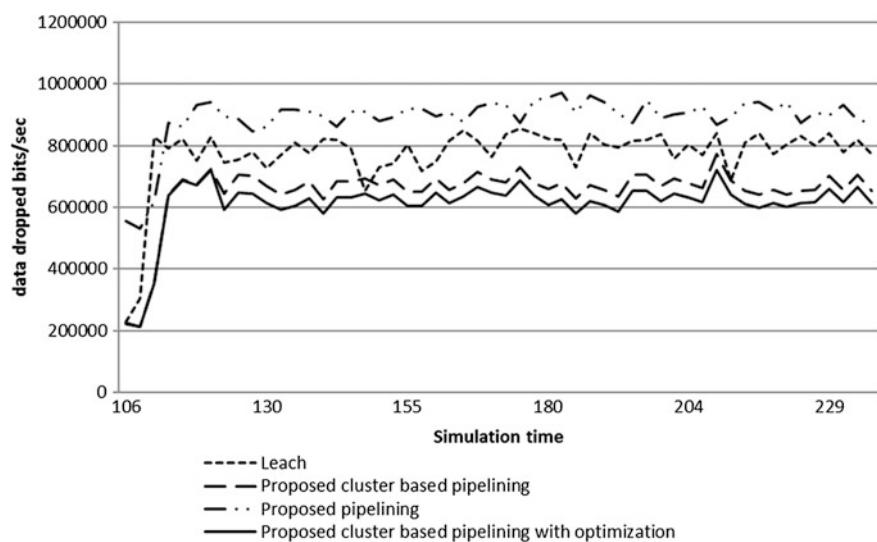


Fig. 6 Data dropped versus simulation time

5 Conclusion

This paper describes about the topic of satisfying throughput, packet delivery time requirement for distributed clustering routing protocol. The issue is a highly suitable topic in the context of the latest development of routing optimizations in WSN. The modified low-energy adaptive clustering hierarchy (LEACH) with BAT routing algorithm improved WSN's routing efficiency. BAT routing techniques can be adopted to work in highly challenging environments. Simulations showed that the proposed BAT optimization performs better in terms of average end-to-end delay and retransmission attempts in comparison with the existing algorithms.

References

1. Lewis, F.L.: Wireless sensor networks. In: Smart Environments: Technologies, Protocols, and Applications, pp. 11–46. Wiley, New York (2004)
2. Kim, C., Koy, Y.B., Vaidya, N.H.: Link-state routing protocol for multi-channel multi-interface wireless networks. In: Military Communications Conference 2008 (MILCOM 2008), pp. 1–7. IEEE, Nov 2008
3. Hu, Y.C., Johnson, D.B., Perrig, A.: SEAD: secure efficient distance vector routing for mobile wireless ad hoc networks. *Ad Hoc Netw.* **1**(1), 175–192 (2003)
4. Ye, W., Heidemann, J., Estrin, D.: An energy-efficient MAC protocol for wireless sensor networks. In: Proceedings of INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies, vol. 3, pp. 1567–1576. IEEE 2002
5. Li, X., Dorvash, S., Chengc, L., Pakzadd, S.: Pipelining in structural health monitoring wireless sensor network. In: Proceedings of SPIE, vol. 7647, pp. 76470I–1, Mar 2010
6. Healy, C.A., Whalley, D.B., Harmon, M.G.: Integrating the timing analysis of pipelining and instruction caching. In: Real-Time Systems Symposium. Proceedings of 16th IEEE, pp. 288–297. IEEE, Dec 1995
7. Liu, X.: A survey on clustering routing protocols in wireless sensor networks. *Sensors* **12**(8), 11113–11153 (2012)
8. Pawa, T.D.S.: Analysis of low energy adaptive clustering hierarchy (LEACH) protocol. Doctoral dissertation (2011)
9. Marriwala, N., Rathee, P.: An approach to increase the wireless sensor network lifetime. In: World Conference on Information and Communication Technologies (WICT), pp. 495–499. IEEE, Oct 2012
10. Farooq, M.O., Dogar, A.B., Shah, G.A.: MR-LEACH: multi-hop routing with low energy adaptive clustering hierarchy. In: Fourth International Conference on Sensor Technologies and Applications (SENSORCOMM), pp. 262–268. IEEE, July 2010
11. Kong, H.Y.: Energy efficient cooperative LEACH protocol for wireless sensor networks. *Commun. Net. J.* **12**(4), 358–365 (2010)
12. Yang, X.S., Karamanoglu, M., Fong, S.: BAT algorithm for topology optimization in microelectronic applications. In: International Conference on Future Generation Communication Technology (FGCT), pp. 150–155. IEEE, Dec 2012
13. Khamfroush, H., Saadat, R., Heshmati, S.: A new tree-based routing algorithm for energy reduction in wireless sensor networks. In: International Conference on Signal Processing Systems, pp. 116–120. IEEE, May 2009

14. Patel, D.K., Patel, M.P., Patel, K.S.: Scalability analysis in wireless sensor network with LEACH routing protocol. In: International Conference on Computer and Management (CAMA), pp. 1–6. IEEE, May 2011
15. Zhang, H., Chen, P., Gong, S.: Weighted spanning tree clustering routing algorithm based on LEACH. In: 2nd International Conference on Future Computer and Communication (ICFCC), vol. 2, pp. V2–223. IEEE, May 2010
16. Li, B., Zhang, X.: Research and improvement of LEACH protocol for wireless sensor network. In: International Conference on Information Engineering (2012)
17. Ran, G., Zhang, H., Gong, S.: Improving on LEACH protocol of wireless sensor networks using fuzzy logic. *J. Inf. Comput. Sci.* **7**(3), 767–775 (2010)
18. Yang, X.S.: A new metaheuristic BAT-inspired algorithm. In: Nature inspired cooperative strategies for optimization (NICSO 2010), pp. 65–74. Springer, Berlin, Heidelberg (2010)
19. Saha, S.K., Kar, R., Mandal, D., Ghoshal, S.P., Mukherjee, V.: A new design method using opposition-based BAT algorithm for IIR system identification problem. *Int. J. Bio-Inspired Comput.* **5**(2), 99–132 (2013)
20. Heinzelman, W.R., Chandrakasan, A., Balakrishnan, H.: Energy-efficient communication protocol for wireless micro sensor networks. In: Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, pp. 10. IEEE, Jan 2000
21. Handy, M.J., Haase, M., Timmermann, D.: Low energy adaptive clustering hierarchy with deterministic cluster-head selection. In: 4th International Workshop on Mobile and Wireless Communications Network, pp. 368–372. IEEE, 2002
22. Abdellah, E., Benalla, S., Hssane, A.B., Hasnaoui, M.L.: Advanced low energy adaptive clustering hierarchy. *Int. J. Comput. Sci. Eng.* **2**(7), 2491–2497 (2010)
23. Khan, K., Sahai, A.: A comparison of BA, GA, PSO, BP and LM for training feed forward neural networks in e-learning context. *Int. J. Intell. Sys. Appl. (IJISA)* **4**(7), 23 (2012)
24. Nakamura, R.Y., Pereira, L.A.M., Costa, K.A., Rodrigues, D., Papa, J.P., Yang, X.S.: BBA: a binary BAT algorithm for feature selection. In: 25th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), pp. 291–297. IEEE, Aug 2012
25. Yang, X.S., He, X.: BAT algorithm: literature review and applications. *Int. J. Bio-Inspired Comput.* **5**(3), 141–149 (2013)
26. Fister, I. Jr, Fister, D., Yang, X.S. A hybrid BAT algorithm. arXiv preprint arXiv:1303.6310 (2013)
27. Taha, A.M., Tang, A.Y.C.: BAT algorithm for rough set attribute reduction. *J. Theor. Appl. Inf. Tech.* **51**(1) (2013)

An Hybrid Ant Routing Algorithm for Reliable Throughput Using MANET

N. Umapathi, N. Ramaraj, D. Balasubramaniam and R. Adlin mano

Abstract The field of wireless networks is an important and challenging area. In this paper, routing in mobile adhoc networks (MANETs) using ant algorithm has been described. ANTHOCNET algorithm makes use of ant-like mobile agents which sample the nodes between source and destination. In MANET, each and every node has an additional task by which it can forward packets between two or more nodes. The routing protocol in MANET should be capable of adjusting between two or more nodes. The routing protocol should be capable of adjusting between high mobility, low bandwidth to low mobility, and high bandwidth scenario. AODV protocol is being estimated for higher throughput. By varying the time cycle, throughput and packet efficiency can be increased. When throughput is increased, efficiency is also high. This also improves the quality of services.

Keywords Mobile adhoc network · ANT colony optimization · AODV · Throughput tolerance

N. Umapathi (✉) · D. Balasubramaniam

Department of Electronics and Communication Engineering, GKM College of Engineering and Technology, Chennai, India
e-mail: nrumpathi@yahoo.co.in

D. Balasubramaniam
e-mail: drdbmaniam@gmail.com

N. Ramaraj
Department of Computer Science Engineering, Thanagavelu Engineering College, Chennai, India
e-mail: ramaraj_gm@yahoo.com

R. Adlin mano
Department of Electronics and Communication Engineering, IFET College of Engineering, Villupuram, India
e-mail: adlinmano@gmail.com

1 Introduction

A mobile adhoc network (MANET) is a collection of nodes communicating with each other without any supporting infrastructure. The routing scheme in a MANET can be classified into two major categories—Proactive and Reactive. The combination of these both protocols is called as hybrid. The need for MANETs is growing, and it can be connected to wired or wireless links using one or more different technologies. In this paper, an application of ant routing algorithm for MANETs is being estimated. The shortest path between the nodes from the source node to the destination node is found. It is mainly used for optimum route discovery in wireless networks.

2 Related Work

(a) *Routing in MANETS*

A family of routing algorithms for MANETS is very popular in different applications considering the following factors: reliability, cost, bandwidth, total required power, ease of installation, security, and performance of network. MANET has several salient features: resource constraints, dynamic topology, no infrastructure, and limited physical security. This routing establishes optimum paths to the destination using a number of artificial ants that communicate with each other by the process called stigmergy. Stigmergy is the process of indirect communication between individuals in an adhoc network.

(b) *ACO Routing*

The information is gathered through path sampling using ant agents. An ant going from source S to destination D collects information about the quality of the path and retraces its way from D to S. This information is used to update the routing tables at intermediate nodes. The routing tables are called as pheromone tables. These pheromone variables are continually updated according to the quality of the paths. The ANTS use pheromone tables to find their way to destination.

ANTS are of two types. Mainly here, the concept of negative backward ants and destination trail ants is being included in order to achieve an optimum solution. This technique reduces the time needed for connection setup and insures the delivery of data. There are many various algorithmic techniques that have been inspired by behavior of ants. ANT colony optimization is one of the most successful techniques and best known among them. Scalable routing is being achieved through AODV protocol. As the total number of nodes in the MANET becomes very large, the overheads of the routing algorithms should be low. It is important to study the various performance metrics for understanding and utilization of routing protocols.

3 Literature Review

AODV is adhoc on demand distance vector routing protocol. It is an improvement of the destination-sequenced distance vector (DSDV) routing protocol. It is based on distance vector which uses destination sequence numbers to determine the kind of the routes. The advantage of AODV is that it tries to minimize the number of broadcasts. It creates routes on on demand basis to maintain a complete list of routes for each destination. The usage of AODV protocol for mobile adhoc networking applications provides consistent results for large scale scenarios.

It is also known as Ford-Fulkerson algorithm. It is a kind of distance vector routing protocols such as RIP, BGP, ISO, IPX, NOVELL, and IDR. This maintains the distance tables which find the shortest path to the sending packets to each node with the help of routers in a network. The information is always updated by exchanging information within the neighboring nodes. The number of data will be equal to the number of nodes in the network. Each data contains the path for sending packets to each destination in a network.

It is a table-driven routing protocol. Each node discovers and maintains topology information in a network. They build the shortest path tree from source to destination. This includes the detection of neighbors and exchanges the topology information among the nodes. When a node receives Hello message from another node, it discovers new neighbor. If a node does not receive any message from the neighbor for a particular period of time, the neighbor is determined which is broken or out of its range.

LAR is location-aided routing protocol. Each node discovers and maintains topology information in a network. They build the shortest path from source to destination. LAR protocol uses GPS to get the location information to reduce the complexity of the protocol. Each host knows its position exactly; the difference between the exact position and calculation position is not being considered.

4 Methodology

Description of Ant Algorithm

ACO algorithms are population-based optimized approach inspired by the behavior of real-life ant colonies. Here, individual ants deposit a substance called pheromone on the path while moving from nest to food sources. Here, individual ants are aided to smell and select the routes. Due to the behavior of ants, the ants would select the shortest path from nest to the food source. The characteristics of artificial ants are as follows:

- Artificial ants are able to memorize their paths.
- Artificial ants are discarded after each time from nest to the food source.
- Artificial ants include heuristic information in this selection.
- Artificial ants live in an environment when time is discrete.

Proposed Ant Algorithm

The proposed ANT algorithm is ANT colony optimization (ACO) algorithm where ANTHOCNET is being used. ANT algorithm requires the following elements:

- Constructions of solutions.
- Pheromone updating table.
- Local search mechanism.
- Heuristic information.
- Termination condition.
- Selection of probability.

Constructions of solutions

Each part of the solution is termed as state. Once a cell is being assigned to a location, the index is taken into a list. The memory is used at the end of the iteration to update the trail levels. Each ant assigns another cell in a location till a complete solution is obtained. This leads each ant to generate a feasible solution.

Pheromone updating table

Real ants deposit a substance called pheromone while moving from one point to another point. The update of each level can be performed either after each move or complete solution.

- Moves with negative trail levels can be avoided. Moves are chosen with equal chances if there is no positive trail level.
- The absolute value is being added to negative trail to become positive. Hence, for negative trail level, there is no chance for selection.

Local search mechanism

ACO heuristic information is used to construct solution. The constructed solution is mainly to update improved solution. Hybrid ant algorithms are included for solving the intercell layout problem.

- Ants are generated randomly based on uniform distribution.
- The distance between the grid points and entries of the flow matrix is created randomly.
- The flow entries must be at most equal to zero and nonzero entries are not uniformly distributed.
- The data problem is generated randomly, and they resemble the distribution of real ants' instances.

Heuristic information

Heuristic information pertaining to the move $v = \pi/\beta$ is denoted by g_i , and the location is calculated using heuristic approach. From which, heuristic information is calculated. From this, the intercell layout problem is also cleared.

Termination condition

ANT algorithm can be terminated in several manners as the algorithm has maximum number of iterations. The value obtained in each iteration is being compared with the earlier iterations.

4.1 Route Discovery

At a node, when a packet is received by the network layer from the higher layers, the node checks to see if routing information is available for destination over any of its neighbors. If found, it forwards the packet over that node, if not, it broadcasts a forward ant to find a path to the destination. When a forward ant is received by an intermediate node, the node checks whether it has routing information for the destination over its neighbors and if found, it unicasts the ant over that neighbor, else it broadcasts the forward ant. Loops are prevented by a sequence id mechanism and endless flooding is restricted by enforcing maximum number of hops for the ant. Once the ant reaches the destination, it becomes a backward ant and it follows the same path it came from, back to the source.

4.2 Route Maintenance

ANTHOCNET uses a proactive updates to improve route quality. Nodes periodically broadcast information about the best pheromone values to each destination at that node. The neighboring nodes on receiving the information then adjust the value of their existing pheromone values of the routing table entries to every destination over the broadcasting node. This diffusion process is slow and could result in new paths being discovered to the destination. However, these paths are not reliable, are thus not used directly in packet forwarding, and are marked as virtual pheromones to be explored later during another route for the ant. Once the ant reaches the destination, it becomes a backward ant and it follows the same path it came from, back to the source.

4.3 Route Error Correction

A link error may be detected when a Hello message is not received from a neighbor for a timeout period, or if a packet fails to transmit though a link. The algorithm

corrects the routing table to reflect the link failure. In the case of packet sending failure, ANTHOCNET checks for alternative routes and if not found initiates a route repair process. It also broadcasts a link failure notification to inform its neighbors about the change in routing information.

4.4 Packet Forwarding

Data packets are forwarded using the probabilities created by FANTs and BANTS, using the colony routing table and local routing table for nodes outside and inside the current node's colony, respectively. The length of the packet queue is considered as an indication of the current level of congestion at the next-hop node, and the next-hop probabilities are adjusted to favor nodes with lower congestion.

5 System Design and Results

For system design of the above system, network load analyses the number of node 25 and varied from 6 to 19 nodes and 400×600 sqm, data payload 1,000 byte per packets, and packet rate 250 k. Network simulator (NS2) has been used to construct the network topology graph. Simulator chosen to evaluate the protocols is NS2. It has important advantages when compared to other simulators. Number of nodes can exceed several thousands. It has more efficient routing tables. It is a binded model between C++ and OTCL. Wireless network performance mainly depends upon throughput and average delay. It is cost-effective of network deployment as wiring is not possible.

The topology is described randomly which is being deployed in a grid. The program is written in C++ and OTCL programming languages. Each link is bidirectional. The graph is generated by trace file under NS2. The increasing in time cycles increases the throughput. The software used here is UBUNTU (Fig. 1).

Here, nodes are established, and routing is done between the nodes from node 19 to node 6. The shortest destination to the nearby node is found. The forward and backward ants checks for the pheromone concentration and establishes the path to the node with higher concentration and transfers the data packets. When the backward ant is received, the packet delays to the node are updated. This is to find the shortest route from a source node to the destination node in the wireless adhoc networks (Fig. 2).

Here, nodes are established, and routing is done between the nodes from node 19 to node 6 then from node 10 to node 4; the shortest destination to the nearby node is found. The forward and backward ants checks for the pheromone concentration and establishes the path to the node with higher concentration and transfers the data packets. When the backward ant is received, the packet delays to the node are updated (Fig. 3).

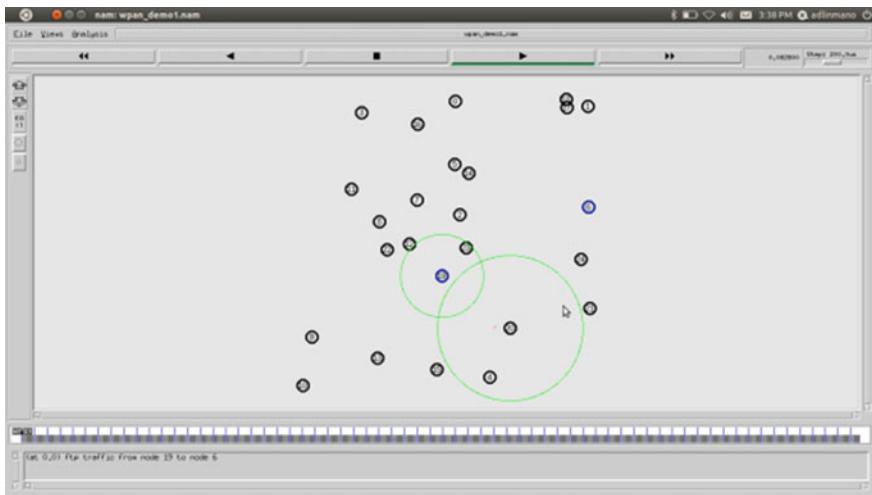


Fig. 1 Network animator from node 19 to node 6

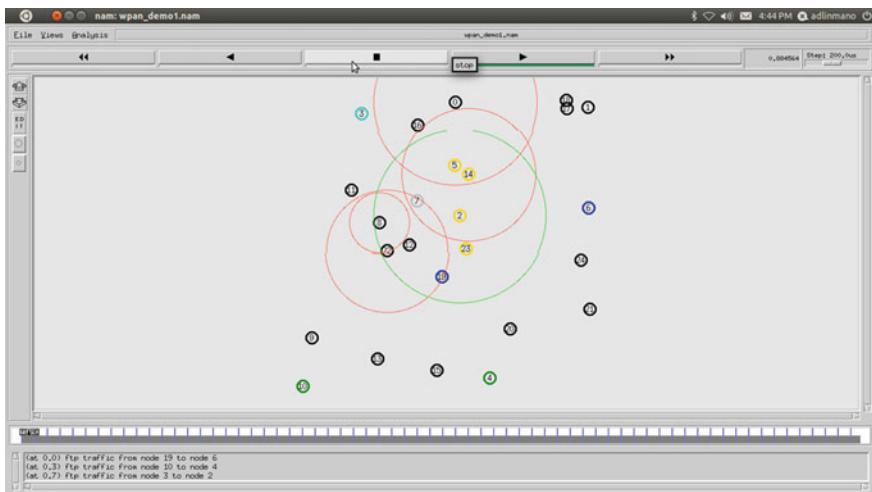


Fig. 2 Network animator from node 10 to node 4

Here, x axis denotes the time cycles, and y axis denotes the throughput of AODV protocol. When time cycles keep on increasing, the throughput data in AODV also increase, thereby improving the packet efficiency. Figure 4 shows time cycles versus AODV throughput. Throughput is being increased by updating the heuristic information after the local search mechanism. The pheromone concentration is being updated with the termination condition given for each iteration. Here, ant

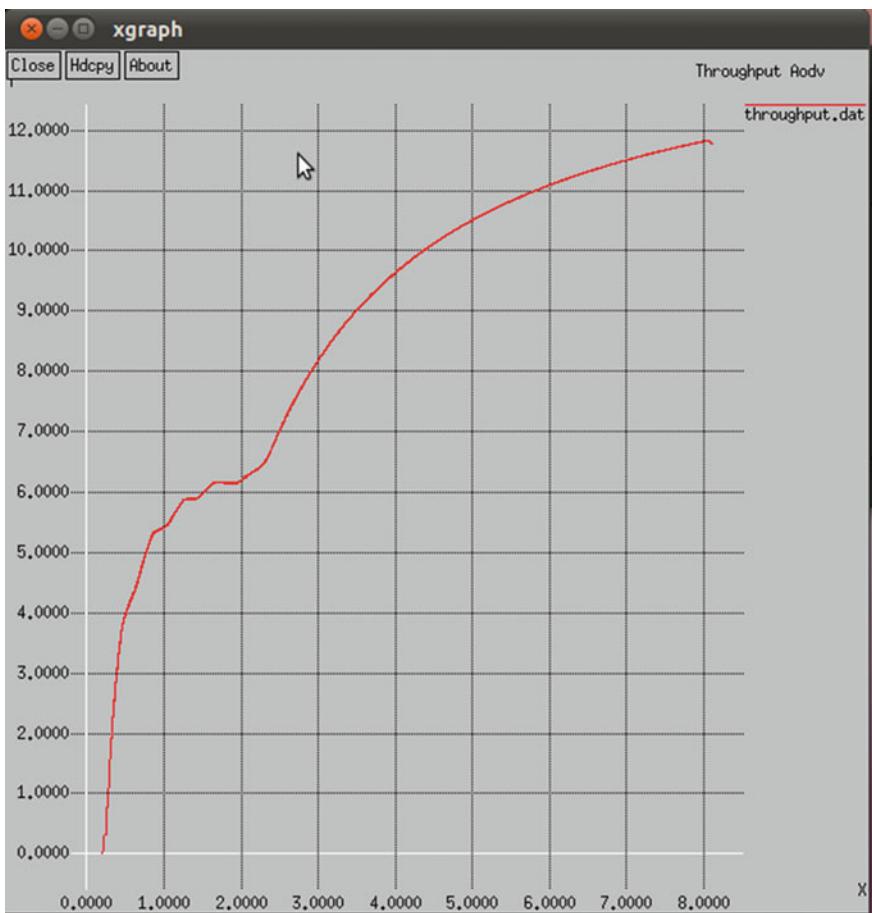


Fig. 3 Time cycles versus AODV throughput

colony termination condition is given for each iteration. Here, ant colony optimization algorithm is being used for delivering high throughput and also for optimization purpose.

Here, x axis denotes the time cycles, and y axis denotes the packet loss. When time cycles keep on increasing, the packet loss for various cycles decreases by which throughput is being increased. Throughput is being increased by improving the congestion window size and optimizing the packet drop rate, thereby providing high efficiency in MANETs.



Fig. 4 Time cycles versus packet loss

6 Conclusion

Hence, the optimization problems are solved by the algorithm, and coding is developed for improving throughput. The shortest path from the source node to the destination node is found, and data packets are being updated. The throughput mechanisms are used to select the forwarding node because velocity does not provide information on link quality. The throughput assists the source node to distribute the forward load to available nodes and avoids the routing holes or deadlock problem. Thus, the optimum throughput is being achieved using AODV protocol.

Bibliography

1. Malany, A.B., Dhulipala, V.R.S., Chandrasekaran, R.M.: Throughput and delay comparison of MANET routing protocols (2009)
2. Sarkar S., Basavaraju, T.G., Puttamadappa, C.: Ad Hoc Mobile Wireless Networks: Principles, protocols and applications, Auerbach Publications, Boca Raton (2008)
3. Tuteja, A., Gujral, A., Thalia, A.: Comparative performance analysis of DSDV, AODV and DSR routing protocols in MANET using NS2. IEEE Comp. Soc. 330–333 (2010)
4. Chakrabarti, S.: Throughput issues in Wireless Adhoc networks (2009)
5. Shah, S., Khandre, A., Shirole, M., Bhole, G.: Performance evaluation of ad hoc routing protocols using NS2 simulation. Mobile Pervasive Comput. 167–171 (2008)
6. Bobade, N.P., Mahala, N.N.: Performance evaluation of adhoc on demand distance vector in MANETs with varying network size using NS-2 simulation. Int. J. Comput. Sci. Eng. 2(8), 2731–2735 (2010)
7. Bhati, P., Chauhan, R., Rathy, R.K.: An efficient agent based AODV routing protocol in MANET. Int. J. Comput. Sci. Eng. 7 (2011)
8. Rajagopalan, S., Shen, C.: ANSI: a swarm intelligence based unicast routing protocol for hybrid ad hoc networks. J. Syst. Archit. 52(8–9), 485–504 (2006)
9. Klein-Berndt, L.: A quick guide to AODV routing. Wireless Communication technology group National Institute of standard and Technology
10. Bouhorma, M., Bentaouit, H., Boudhir, A.: Performance comparison of ad-hoc routing protocols AODV and DSR. Int. Conf. Multimedia Comput. Syst (2009)

Development of Type 2 Fuzzy Rough Ontology-based Middleware for Context Processing in Ambient Smart Environment

A.B. Karthick Anand Babu and R. Sivakumar

Abstract There is currently lot of work in ambient intelligence particularly in context awareness. Context awareness enables service discovery and adaptation of computing devices for ambient intelligence application. There is a common agreement of the fact that context-aware systems should be responsive to multi-agents, covering a large number of devices, assisting a large number of people and serving a large number of purposes. In an attempt to achieve such context-aware systems with scalable scenario implementations, we propose an adaptive and autonomous context-aware middleware for multi-agents with Type 2 fuzzy rough context ontology. Our model provides a meta-model for context description that includes context collection, context processing and applications reactions to significant context changes.

Keywords Ambient intelligence · Type 2 fuzzy rough context ontology · Context awareness

1 Introduction

The recent advances in ubiquitous computing, ubiquitous communication and intelligent user interfaces make ambient intelligence more dynamic and proactive vision of the futuristic computing. Context awareness enables service discovery and adaptation for ubiquitous computing devices. The term ‘context awareness’ was first explicitly introduced in the research area of pervasive computing in it refers to the ability of computing systems to acquire and reason about the context information

A.B. Karthick Anand Babu (✉) · R. Sivakumar
Department of Computer Science, A.V.V.M. Sri Pushpam College (Autonomous),
Poondi, Thanjavur, Tamil Nadu, India
e-mail: dta_babu@yahoo.co.in

R. Sivakumar
e-mail: rskumar.avvmspc@gmail.com

and adapts the corresponding applications accordingly. Context is a key issue for many research communities, such as ambient intelligence, real-time systems or mobile computing, because it relates information processing and communication to aspects of the situations, in which such processes occur. A system that recognizes the situation automatically using various sensors is termed as context-aware system. Therefore, different user devices need semantically rich descriptive context models to provide shared understanding. Ontology modelling is one such context modelling technique used as specific formal specifications in an exceedingly domain and also the relations among them. Ontology modelling is widely accepted tool for modelling context information in pervasive computing domain. In context-aware computing, knowledge representation to user is vague or uncertain which harms users' confidence in using context-aware applications. This kind of system will lead to insecure knowledge representation. Moreover, it is well known that many real-time context-aware applications require support for managing this imprecise vague context to avoid inconsistency and integrity. Intelligent solutions are needed to cope up with the fuzziness of context information and, especially, because of mobility, rapidly changing environments and unsteady information sources. Several methods have been proposed and applied to deal with vague contextual information [2]. The notion of uncertainty or vague knowledge representation to user in context-aware computing appears mainly as a consequences of the complexity of context acquisition, context processing and context representation. To handle this uncertainty, we proposed a context-aware middleware to deal with vague knowledge through **Type 2 fuzzy rough** context ontology rules. Efficient rules in context-aware computing can be used both for improving the quality of context information as well as for deriving higher level probabilistic contexts. The rules can also be used for resolving conflicts between context information obtained from different sources. Such rules not only cope with the imprecise knowledge, but also with the user behaviour and his/her historical context. In this paper, we also perform the adaptation process by using **Type 2 fuzzy rough** to define user's context and rules for adopting the policies of implementing a service through effective rule acquisition.

Sections 2 and 3 discuss related works and the main objectives of the proposed work. Sections 4 and 5 discuss the architecture, methodology with a case study to demonstrate a context-aware application. Finally, Sect. 6 concludes the paper with future discussion.

2 Related Work

Hongyuan et al. present a rule-based context-aware adaptation model in which a rich landscape in adaptation is specified, but a complete and generic context-aware adaptation approach is missing. Several surveys dealt with service composition. Many of them classified the middleware under exclusive criteria such as manual versus automated and static versus dynamic. Agostini et al. [5] illustrate a hybrid approach where ontological reasoning is loosely coupled with the efficient rule-based

reasoning of middleware architecture for service adaptation. Zarandi et al. [8] developed a type 2 fuzzy rule-based expert system for stock price analysis. Interval type 2 fuzzy logic system permits to model rule uncertainties, and every membership value of an element is interval itself. Wang et al. implemented a self-guided ontology-based mobile data management model for the context information. Lemos et al. present an automated redirection of incoming communications in RING, a rule-based context-aware recommender system using semantic web. RING lets users to receive any quite communication through the most effective channel accessible on personal preferences and context. Giang et al. deal with uncertainty using fuzzy logic, but they do not concentrate on exact service to user context. Service ought to consider uncertainty represented in their models. We argue that fuzziness on context information must be considered when selecting services. In the literature, many platforms were proposed to facilitate adapting context-aware services.

3 Objectives

This paper presents *Type 2 fuzzy rough* ontology-based approach for developing context-aware middleware for the following objectives: (1) a well-defined, unified ontology enables knowledge sharing and reuse with declarative semantics which provide multiple policies to support context inference. (2) Ontologies provide complex efficient inference mechanisms to deduce high-level contexts from low-level, raw context data and to check inconsistent contextual information due to imperfect sensing. (3) Explicitly represented ontologies enhance the development of context-aware systems with semantic web technologies.

4 Architecture

The proposed context-aware middleware architecture, showed in Fig. 1, provides an ontology-based model for creating context-aware applications. The proposed model constitutes five layers: at the lowest layer, we have access to different computational entities and diverse sensors and actuators devices. Next layer is an interface allowing two ways exchange between devices and middleware management layer.

Application Layer			
User Context Aware Security Layer			
Reasoning Engine	Context Knowledge Base	Context Query Engine	Context Data base
Sensor Data Fusion and Secure Information interoperability Layer			
Sensor Data Agents for Smart phone, wearable computer, and tablet			

Fig. 1 Context-aware middleware architecture

Third layer is autonomous middleware management layer which contains an adaptive reasoning engine, a context knowledge base (KB), a context database and a context query engine. Next layer is a high-level user context-aware security interface between the middleware and the application layer.

The proposed middleware is a component-oriented middleware which automates context management by adding context-aware components to the middleware and uses container characteristics to automate adaptation process as it is described in this section. Adaptation to context changes is achieved by analysing and interpreting context data using Type 2 fuzzy rough ontology.

5 Methodology

5.1 Design Considerations

The implementation of context awareness is made possible by a mixture of divergent technologies. First, the system needs appropriate data collection technologies for sensors and fusing technologies at a higher level for analysing and interpreting sensor data. Decoupling the sensing layer from applications is the key principle design for context-aware applications. Software frameworks are needed to maximize the usefulness of context information for different services, and designers need to create widely acceptable model with standard mechanisms together for accessing these context information for an effective context-aware system.

5.2 Context Modelling

5.2.1 Ontology-based Context Modelling Approach

The main advantage of ontology context model is sharing of context data among users for common better cognition. This enables devices and services to enable semantic interoperability to provide services to the individual user. This model enables formal analysis and reuse of domain knowledge. The use of ontology will make the application environment independent of programming. The proposed ontology model is divided into three levels as shown in Fig. 2. Context can be obtained from the source provider directly and/or obtained by integration and

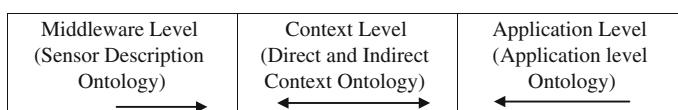


Fig. 2 The proposed model

reasoning of direct context. The first ontology is associated to the middleware level, and it concerns sensors description. This ontology contains information about sensors that the middleware can interact with. This information is created and updated easily by the middleware maintenance agent to enable the middleware to interact with new sensors. Sensors can be used by all applications running on top of this middleware.

The second ontology of Fig. 2 is associated to the context level. It gathers information from sensors to provide direct and indirect context data to context-aware applications. The direct context is captured directly from sensors, and the indirect context is interpreted from other contexts. It also supports interpretation rules and adaptation policies. Two types of adaptation are taken into account: reactive and proactive adaptations. Reactive adaptation consists in listening to changes in the execution context and reacting accordingly, whereas proactive adaptation, which concerns service request, consists in anticipating the way in which the service is provided by changing the source code of the service policy by another policy according to context changes. To describe context classification information in our context, we introduced property element—*owl: classifiedAs* in property restrictions. This element can capture related properties about data types and object in the context classification. The values of the property element are ***aggregated, deduced, sensed or defined***. After determining classification of property, considering the interdependence between context information, we further analyse the characteristics of context information. To describe restraint in context information by *OWL*, we introduced property *element*—*rdfs: dependsOn* in object properties and data attributes—which is to capture attribute dependence between data types and object. The third ontology is associated with the application level, and it gathers information specific to the application allowing the designer to describe using *RelevantContext* class.

5.2.2 Example Scenario

To illustrate the use of the proposed model for the creation of context-aware applications, we give in this section a simple example for laboratory room management. This room is equipped with indoor RFID sensors. Each student and faculty wears an RFID tag; this tag provides data to detect the role of participant. RFID readers are placed in the laboratory room to enable capturing signals emitted by each RFID tags. When a person enters in the laboratory room, presence is detected, and then, the related lab exercises and assignments are automatically transferred to the laptop or PDA. The laboratory room management application is made up of three components. The *Labsession* component is installed in the server, and it contains all lab sessions in a semester. The *Labdetails* component is a plug-in that supports divergent devices of students and faculty members; it enables to download the laboratory exercise programs and all the related information for understanding the subject. The *ViewAgent* Component is a plug-in divergent devices which enables to display a selected laboratory details in the user device.

The laboratory room application is sensitive about *StudentUnderstanding* which is captured directly from the RFID sensor and *ProximityToSubjectUnderstanding* which is interpreted from *StudentUnderstanding*. RFID sensor allows, the sensor instance (RFID) described in the ontology allows the middleware to interact with the master server, in order to give the details of person. The interpretation rules are used to interpret *ProximityToStudentUnderstanding*, and this context is deduced by comparing the student current understanding with the subject grade level using the *NearUnderstanding* method. The *NearUnderstanding* method as instance of the *Skill&Grade* determines whether a student understanding to the subject is clear in the laboratory with its *Skill* and *grade*.

6 Conclusion

Context awareness helps applications running on mobile devices to react into changing environment conditions. An efficient context representation and a context-aware middleware infrastructure are necessary for helping developers to create such type of applications. In this paper, we proposed a context-aware middleware which provides ontology model for context description and application adaptation. The proposed model allows context and relevant context are to be interpreted in the middleware for relevant context adaptation and delivery. Comparing to the existing context models solutions, the proposed model adds the possibility to describe actions of adaptation and sensors from which data are collected, and the relationships between different classes of the ontology are predefined by the proposed model. These relationships allow the middleware to manage effectively the collected context information and to adapt applications and services to context changes.

References

1. Wang, H., Mehta, R., Chung, L., Supakkul, S., Huang, L.: Rule-based context-aware adaptation: a goal-oriented approach. *Int. J. Pervasive Comput. Commun.* **8**(3), 279–299 (2012)
2. Baader, F., Kusters, R., Molitor, R.: Rewriting concepts using terminologies. In: Cohn, G., Giunchiglia, F., Selman, B. (eds.) 7th International Conference on Principles of Knowledge Representation and Reasoning (KR2000), pp: 297–308 (2000)
3. Wang, B., Liu, D., Wong, S.: A context information ontology hierarchymodel for tourism-oriented mobile E-commerce, *J. Soft.* (1796217X) **7**(8), 1751–1758 (2012)
4. Lemos, M.L., Vasquez, D.V., Radzimski, M., Lemos, A.L., Gómez-Berbís, J.M.: RING: a context ontology for communication channel rule-based recommender system. *E meeting of the Proceedings of SeRSy* (2012)
5. Agostini, A., Bettini, C., Riboni, D.: Loosely coupling ontological reasoning with an efficient middleware for context-awareness. In: *Mobiuitous*, The Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services, pp. 175–182 (2005)

6. Jiang, Y., Dong, H.: Uncertain context modeling of dimensional ontology using fuzzy subset theory. In: Proceedings of the 2nd International Conference on Scalable Uncertainty Management, pp. 256–269 (2008)
7. John, R., Coupland, S.: Type-2 fuzzy logic—a historical view. *IEEE Comput. Intell. Mag.* **2**, 57–62 (2007)
8. Zarandi, M.H.F., Neshar, E., Turksen, I.B., Rezaee, B.: A type-2 fuzzy model for stock market analysis. *Fuzzy Systems Conference, 2007. FUZZ-IEEE 2007. IEEE International*, pp. 1–6, 23–26 (2007). doi:[10.1109/FUZZY.2007.4295378](https://doi.org/10.1109/FUZZY.2007.4295378)
9. Yang, S.J.H., Zhang, J., Chen, I.Y.L.: A JESS-enabled context elicitation system for providing context-aware web services. *Expert Syst. Appl.* **34**, 2254–2266 (2008)
10. Schilit, W.N.: A system architecture for context-aware mobile computing. Ph.D. thesis, Columbia University, New York 1(995)
11. Strang, T., Linhoff-Popien, C.: A context modeling survey. In: 1st Int'l Workshop on Advanced Context Modelling, Reasoning and Management. pp. 34–41 (2004)
12. Gruber, T.R.: A translation approach to portable ontology specification. *Knowl. Acquisition* **5** (2), 199–220 (1993)

Multiple-Image Encryption Using Genetic Algorithm

Subhajit Das, Satyendra Nath Mandal and Nabin Ghoshal

Abstract Multiple images have to be sent some time for different purposes. Those images encrypt individually before their sending through non-secure channel. This process is painful and also time-consuming. In this paper, multiple images are encrypted using genetic algorithm. Here, a set of crossover points and a set of values are used to diffuse and encrypt the images. These two sets are computed from a key sequence. The original images have been diffused by genetic algorithm. Bitwise XOR operation has been applied between key set and diffuse images to get encrypted images. The effectiveness of the algorithm has been tested by number of statistical tests like histogram analysis, correlation, and entropy test. A comparison is made between the proposed algorithm and other genetic-based encryption algorithm. Finally, it has been observed that the proposed algorithm is given better result in all tests with less execution time.

Keywords Multiple image · Genetic algorithm · Histogram analysis · Correlation and entropy test

1 Introduction

For security purpose, image data have been encrypted when they are sending through non-secure channel. Different algorithms apply on image data, and they encrypt image according to the direction of algorithms. But, traditional encryption

S. Das (✉)
Nayagram Bani Bidyapith, Paschim Medinipur, India
e-mail: subhajit.batom@gmail.com

S. Mandal
Department of Information Technology, Kalyani Government Engineering College,
Kalyani, Nadia, India
e-mail: satyen_kgec@rediffmail.com

N. Ghoshal
Department of Engineering and Technology Studies, University of Kalyani, Nadia, India
e-mail: nabin_ghoshal@yahoo.co.in

algorithms are not very effective for image encryption because of larger amount of data, higher redundancy, and stronger correlation between pixels [1]. Zeghid et al. [2] have developed an image encryption method based on modified AES (Advanced Encryption Standard). Different authors have also used Blowfish and Hill cipher algorithms in different ways for image encryption [3, 4]. Yen and Guo [5] have proposed Chaotic Key-Based Algorithm (CKBA) for image encryption. Their proposed algorithm relies on a one-dimensional chaotic map for generating a pseudorandom key sequence. Socek et al. [6] have enhanced the CKBA algorithm by replacing the one-dimensional chaotic logistic map to a piecewise linear chaotic map to improve the performance of image encryption algorithm. In recent years, genetic algorithm-based image encryption has found wide popularity among researchers. Al-Husainy [7] have demonstrated a symmetric key image encryption using genetic algorithm. Soni et al. [8] described a key generation algorithm based on genetic algorithm. But, most of the authors have used different image encryption techniques on a single image to produce cipher image. The encrypted images are also forming some pattern as huge amount of similarities present in neighbor pixels. In this paper, multiple images are encrypted at a time. The original images have been diffused by genetic algorithm to eliminate the neighborhood pixels similarities. A bit sequence that is multiplicative of 64 has been taken, and it is divided into equal parts so that each part contains 16 bits. The prime numbers are computed from each 16 bits numbers and placed them back to back to get the effective key and crossover point sets. It is observed that time complexity of this approach is reduced compared to other algorithms. This type of work has never been done in past. This is the reason for making this paper.

2 Proposed Algorithm

Our proposed method to encrypt two grayscale images has been divided into five parts: key generation, image diffusion, encryption, decryption, and image reconstruction. In key generation method, two effective key sets have been generated by a set of bit stream (multiplicative 64) and applied on remaining four steps of the algorithm. The image pixels are altered by genetic algorithm to produce better encrypted image. Logical operation has been applied between diffused images and key set to produce encrypted images.

2.1 Key Generation

Input: Key stream (multiplicative of 64).

Output: Two set of remainder number.

For demonstration, it has been assumed that 64-bit key stream has been taken, but for better key space, we should take 128-bit key stream for our algorithm.

Method:

Step 1: A sequence of 64 bits (\mathbb{K}) has been taken as initial key value.

Step 2: To divide \mathbb{K} into 4 parts where each part consists of 16 bits.

$$\mathbb{K} = \{K_i | 1 \leq i \leq 4\}$$

Step 3: To compute decimal value of each K_i .

$$D_i = \text{decimal}(K_i) | 1 \leq i \leq 4$$

Step 4: To compute all the prime numbers ranging from 2 to D_i and stored in D_i^{pr} into reverse order.

$$D_i^{\text{pr}} = \text{rev}(\text{primes}(D_i)) | 1 \leq i \leq 4$$

Step 5: To store all the D_i^{pr} one by one and store them in array.

$$E = \{D_1^{\text{pr}} + D_2^{\text{pr}} + D_3^{\text{pr}} + D_4^{\text{pr}}\} | + \text{ denotes concatinaion of arrays.}$$

Step 6: To divide all the numbers by 256 and 8 and store their remainder into E_k and E_p , respectively.

$$E_k = \text{remainder}(E, 256), \quad E_p = \text{remainder}(E, 8)$$

2.2 Image Diffusion

Input: Two grayscale images, set of remainder E_p .

Output: Two diffused image.

Method:

In this method, crossover and mutation operations (complementing the LSB of each pixel) of genetic algorithm have been performed between the pixels of two images.

Step1: Two gray scale image is a set of pixels represented by matrix \mathbb{M} and \mathbb{N}

$$\mathbb{M} = \{a_{ij}\}_{m \times n} \text{ and } \mathbb{N} = \{b_{ij}\}_{m \times n}.$$

Eight bit binary equivalent of the elements of \mathbb{M} and \mathbb{N} can be represented as

$$a_{ij} = a_1^{ij} a_2^{ij} \dots a_8^{ij} \text{ and } b_{ij} = b_1^{ij} b_2^{ij} \dots b_8^{ij}$$

Step 2 : set $count = 1$

for $i = 1$ to m

for $j = 1$ to n

$r = E_p(count);$

$$C1^{ij} = \{a_1^{ij} a_2^{ij} a_r^{ij} b_{r+1}^{ij} b_{r+2}^{ij} \dots \sim b_8^{ij}\} \&$$

$$C2^{ij} = \{b_1^{ij} b_2^{ij} b_r^{ij} a_{r+1}^{ij} a_{r+2}^{ij} \dots \sim a_8^{ij}\}$$

(where $1 \leq r \leq 8$ and \sim represent complement operation)

$count = count + 1;$

End of j

If ($count > \text{size}(E_p)$) then $count = 1$;

End of i .

2.3 Image Encryption

Input: Two diffused images and set of remainder E_k .

Output: Two cipher image.

Method:

Logical XOR operation is performed on diffused image matrixes $C1, C2$ element by element with the elements set E_k .

$$C1' = \{C1 \oplus E_k\} \text{ and } C2' = \{C2 \oplus E_k\}$$

If total number of elements of diffused image is greater than elements of remainder set E_k , then it has been taken in a circular manner.

2.4 Image Decryption

Input: Two cipher images and set of remainder E_k .

Output: Two diffused images.

Method:

Like encryption, same operation has been performed on cipher images to obtain two diffused image $C1$ and $C2$. i.e.

$$C1 = \{C1' \oplus E_k\} \quad \text{and} \quad C2 = \{C2' \oplus E_k\}$$

2.5 Image Reconstruction

Input: Two diffused images and set of remainder E_p .

Output: Two input image.

Method:

```

set count =1
for i= 1 to m
    for j= 1 to n
        r =  $E_p(count)$ ;
         $a^{ij} = \{ c1_1^{ij} \ c1_2^{ij} \ c1_r^{ij} \ c2_{r+1}^{ij} \ c2_{r+2}^{ij} \dots \dots \sim c2_8^{ij} \} \ &$ 
         $b^{ij} = \{ c2_1^{ij} \ c2_2^{ij} \ c2_r^{ij} \ c1_{r+1}^{ij} \ c1_{r+2}^{ij} \dots \dots \sim c1_8^{ij} \}$ 
        ( where  $1 \leq r \leq 8$ , and  $\sim$  represent complement operation)
        count = count + 1;
    End of j
If (count >size(  $E_p$ )) then count=1;
End of i.

```

3 Experimental Setup

A PC with Intel Core 2 Duo 1.50 GHz CPU, 1 GB RAM, 500 GB hard disk with Windows7 operating system and with MATLAB 7.9.0.529 have been used to perform the proposed method. Gray-level images with dimension 256×256 and 512×512 have been taken for experiments.

4 Experimental Result

4.1 Key Sensitivity Test

According to the principal of key sensitivity, it is said that slight change in key value never produce original image in decryption. In this test, a 128-bit key stream



Fig. 1 **a** Input images. **b** Decrypted images with right key. **c** Decrypted images with wrong key

has been applied on lena and babun images. The encrypted images are successfully decrypted by same key value. Now, one bit in key value has been changed and the new key is applied on same encrypted images. It has been observed that it produced completely different images. The result of the above experiments has been furnished in Fig. 1. It proves the key sensitivity of proposed method.

4.2 Statistical Analysis: Histogram Analysis and Correlation Analysis

Histogram for images is a graph where the X -ordinate represents the gray-level value and the Y -ordinate represents the occurrence frequency for each gray-level value. An encryption quality is said to be good if histogram of the cipher image is almost uniform in Y values [9, 10]. It is very difficult for an attacker to extract plain image from encrypted image after analyzing its statistical nature. Histograms of different plain image and cipher images have been presented in Table 1. From Table 1, it is clear that it is quite impossible to get original images from encrypted images.

An encryption technique is said to be good if the technique breaks the correlation among the adjacent pixels of the image. For this purpose, the correlation between two adjacent pixels is calculated and if it is near about 1, it is said a good encryption [11]. In this experiment, randomly 1,000 pairs of adjacent pixels (either in horizontally or vertically or diagonally) have been selected and then calculate their correlation using Eqs. 1–4.

$$E(x) = \frac{1}{N} \sum_{i=1}^N x_i \quad (1)$$

$$D(x) = \frac{1}{N} \sum_{i=1}^N (x_i - E(x))^2 \quad (2)$$

$$\text{COV}(x,y) = \frac{1}{N} \sum_{i=1}^N (x_i - E(x))(y_i - E(y)) \quad (3)$$

Table 1 Experimental result

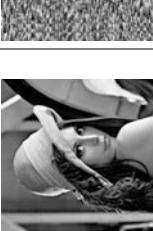
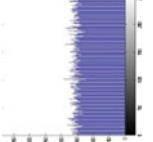
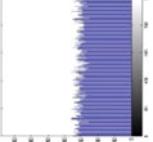
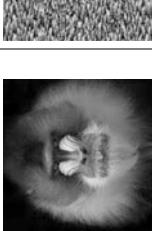
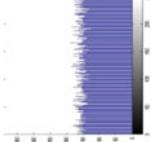
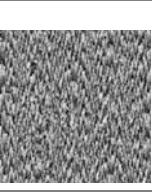
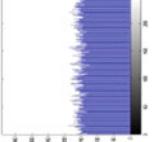
Experiment no.	Input_image	Cipher_image	Histogram cipher image	Correlation coefficient			Entropy	Execution time (s)
				Vertical	Horizontal	Diagonal		
Experiment 1				0.0061	0.007	0.0084	7.9961	0.34843
				0.0084	0.0032	0.0102	7.9959	
Experiment 2				0.0048	0.0076	0.0051	7.9956	0.32881
				0.0054	0.0047	0.0053	7.9959	

Table 2 Comparison with different algorithm based on genetic algorithm

Input image	Methods	Correlation coefficient			Entropy	Execution time (s) (key generation + diffusion + encryption)
		Vertical	Horizontal	Diagonal		
Babun	Proposed method	0.0048	0.0076	0.0051	7.9956	0.32881
Monalisa		0.0054	0.0047	0.0053	7.9959	
Babun	Ref. [12]	-0.0051	0.0085	-0.0023	7.9933	0.4286
Monalisa		0.0012	0.0035	0.0029	7.995	0.3965

$$r = \frac{\text{COV}(x, y)}{\sqrt{D(x)} \sqrt{D(y)}} \quad (4)$$

where x and y are grayscale values of two adjacent pixels in the image. The correlation coefficients of horizontal, vertical, and diagonal pixels of input and cipher images are shown in Table 1. This correlation analysis shows that the proposed encryption algorithm is satisfying zero co-correlation.

4.3 Entropy

Entropy is one of the prominent features in randomization. Entropy of an image [6] is calculated as

$$H(m) = \sum p(m_i) \log 2 \frac{1}{p(m_i)} \quad (5)$$

where $p(m_i)$ represents the probability of the pixel value m_i . Entropy value of different cipher images is also furnished Table 1. It is observed that all entropy values are near 8.

4.4 Performance Comparison with the Ref. [12]

The proposed algorithm has been compared with the previous algorithm (Ref. [12]). The performance of two algorithms has been furnished in Table 2.

5 Conclusion

In this paper, genetic algorithm-based multi-image encryption technique has been discussed. It has been observed that the algorithm is given better result in different statistical tests. The key sensitivity test is shown that it has been given different

result after changing one bit in key value. The histogram analysis is shown that gray level of the distributed pixels in encrypted images varying from 0 to 255 at almost same frequencies. The entropy values are almost near equal to 8 indicated the good encryption quality. Finally, the time complexity is much less than other genetic algorithm-based encryption applicable in single image. The other statistical tests and the algorithm will be applied on more images in future.

References

1. Wang, X., Luan, D.: A novel image encryption algorithm using chaos and reversible cellular automata. *Commun. Nonlinear Sci. Numer. Simul.* **1**, 3075–3085 (2013)
2. Zeghid, M., Machhout, M., Khrijji, L., Baganne, A., Tourki, R.: A modified AES based algorithm for image encryption. *Int. J. Comput. Inf. Syst. Control Eng.* **1**(3), 735–740 (2007)
3. Younes, M.A., Jantan, A.: Image encryption using block based transformation algorithm. *Int. J. Comput. Sci.* **35**, 1 (2008) (9 pages)
4. Yadav, R.S., Beg, R., Tripathi, M.M.: Image encryption technique: a critical comparison. *Int. J. Comput. Sci. Eng. Inf. Technol. Res. (IJCSEITR)* **3**(1), 67–74 (2013)
5. Yen, J.-C., Guo, J.-I.: A new chaotic key-based design for image encryption and decryption. *IEEE Int. Conf. Circuits Syst.* **4**, 49–52 (2000)
6. Socek, D., Li, S., Magliveras, S.S., Furht, B.: Enhanced 1-D chaotic key-based algorithm for image encryption. In: *IEEE/CreateNet SecureComm*, pp. 406–408. Athens, Greece, 5–9 Sept 2005
7. Al-Husainy, M.A.F.: Image encryption using genetic algorithm. *Inf. Technol. J.* **5**, 516–519 (2006) (Asian Network For Scientific Information)
8. Soni, A., Agrawal, S.: Using genetic algorithm for symmetric key generation in image encryption. *Int. J. Adv. Res. Comput. Eng. Technol. (IJARCET)* **1**(10), 137–140 (2012). ISSN: 2278–1323
9. Guanghui, C., Kai, H., Yizhi, Z., Jun, Z., Xing, Z.: Chaotic image encryption based on running-key related to plaintext. *Sci. World J. Article id 490179*, p. 9 (2014)
10. Bhatnagar, G., JonathanWu, Q.M.: A novel chaotic encryption framework for securing palmprint data. In: *The 3rd International Conference on Ambient Systems, Networks and Technologies*, pp. 442–449. Elsevier (2012)
11. Jolfaei, A., Mirghadri, A.: A novel image encryption scheme using pixel shuffler and A5/1. In: *International Conference on Artificial Intelligence and Computational Intelligence*, pp. 369–373 (2010)
12. Das, S., Mandal, S.N., Ghoshal, N.: Diffusion of encryption of digital image using genetic algorithm. *FICTA 2014 Advances in Intelligent Systems and Computing*, vol. 327, pp. 729–736. Springer (2015)

Knowledge-Based System Architecture on CBR for Detection of Cholera Disease

**Souvik Chakraborty, Chiranjit Pal, Shambo Chatterjee,
Baisakhi Chakraborty and Nabin Ghoshal**

Abstract Case-based reasoning (CBR) is an appropriate methodology that applies logical reasoning using similarity measure to relate a current problem case with past similar cases. It has been applied successfully in medical diagnosis and has been experimented in different domains of application in diagnosis and detection. In this paper, we have proposed knowledge-based decision support system which uses the concept of CBR to detect cholera disease. CBR is problem solving method which is derived from artificial intelligence and is based on some base cases which can be revised in order to determine homogeneous cases for new problem. Experimental results show that the proposed model Cholera Easy Detection System (CEDS) assists the doctors to make a consistent decision. Through this work, we are intending to provide facility to the medical research scholars as well as medical unit in order to help them identify cholera when the patient is infected with correspondence symptoms of that disease. Moreover, the CEDS also assists in minimizing errors of deviation that have been found to be noticeable cause of medical errors.

Keywords Case-based reasoning • Detection • Cholera • Knowledge-based systems • Medical diagnosis and detection

S. Chakraborty (✉)
Hooghly Engineering and Technology College, Hooghly, India
e-mail: souvik.india94@gmail.com

C. Pal · N. Ghoshal
Department of Engineering and Technological Studies, University of Kalyani, Kalyani, India
e-mail: chiranjit.pal.14@gmail.com

N. Ghoshal
e-mail: nabin_ghoshal@yahoo.co.in

S. Chatterjee
Kalyani Mahavidyalay, Kalyani, India
e-mail: shambo.india94@gmail.com

B. Chakraborty
National Institute of Technology, Durgapur, India
e-mail: baisakhichak@yahoo.co.in

1 Introduction

The aim of our project is to explore the principle of CBR in developing Cholera Easy Detection System (CEDS). Case-based reasoning (CBR) is a technique presented by Riesbeck and Schank [1] and Kolodner [2] based on human behavior. Aamodt and Plaza [3] introduce a basic model for developing CBR applications. CBR is said to be a technique that deals with problems by remembering similar formal cases which are stored in case base and comparing present cases with them [4] just how a human brain works. Generally, the life cycle of CBR is described by 4 phases, such as retain the most similar cases to compare with new problem, reuse the cases to attempt solution of the problem, revise the proposed solution if necessary, and obtain the solution as part of the new case. So CBR can also be represented as the knowledge-based learning approach as each time a new case occurs, the system memorizes the case.

In this work, we have proposed a model CEDS that detects Cholera disease through some certain symptoms felt by the patient. In our proposed approach, the symptoms are classified as primary and secondary symptoms based on their intensity. The system diagnoses the problem in two stages. In first stage, the system uses cardinality approach, and in the second stage, it uses similarity searching approach.

Cardinality approach: In this approach, system will check number of primary symptoms experienced by the patient he or she has related to the disease to generate an initial solution.

Similarity searching: In this approach, system will determine the similarity of formal cases with the new case to select the most similar case based on some mathematical calculations as explained in Sect. 4.

Our proposed system assists medical experts to take an immediate action by returning an automated solution on the basis of some input parameters. This work is an empirical study of CBR in cholera disease detection. In future work, we like to explore the proposed system with more real-time cases.

2 Literature Study

CBR has played a significant role in modern technology and has widely been used in the field of medicine as in [5], aircraft conflict resolution as in [6], and military decision support system as in [7] and several other fields.

The working principle of CBR mainly depends on two main parts as in [3]: The first part is the retrieval, in which most historical cases are retained from case base to find the similarity between old case and the new one. Depending on the size of case base, many algorithm are deployed, namely nearest neighbor match and ID3 as in [8]. The second part is the adaptation which suggests solving the problem by learning from experience. It means solutions of past similar cases have to be

updated if required to fit for a new case. Since adaptation is a very complex method and field oriented, no generalized adaptation methods have been developed so far.

In [9], a knowledge-based viral fever detection system has been developed which is based on CBR. This system is based on viral fever cases and symptoms by which one can detect the type of viral fever he or she has and what are the necessary action he or she has to take. A patient may report with some symptoms of different severity level (low, medium, and high) corresponding to each symptom. From the value of the severity of the symptom in a patient and the weight of each symptom with respect to a disease, the CBR system calculates the weighted mean value of each disease by which VFDS will detect the fever type that has the closest similarity with the existing cases in the case base.

The paper [10] discusses a knowledge-based system that uses rule-based and case-based method together to achieve the diagnosis. Rule-based systems deal with the problems with well-defined knowledge bases which do not provide flexibility of such system. CBR is used to extend the features of rule-based systems by utilizing the formal cases with new cases to improve the performance of the system. The result of this research shows that the system is useful to help pathologists in making accurate and timely diagnoses. The system also helps in eliminating errors which are a prominent cause of medical errors. The input deals with some specific characteristics corresponding to the disease, and the output is the probable solution to that disease. The system updates its case base based on the differences in characteristics between the old and new cases.

The research in [11] has shown the different AI approach for solving complex medical diagnosis in the determination of cancer which claims to be a better diagnosis method than the conventional human diagnosis which is significantly worse than the neural diagnostic system. This paper describes a new system for detection of heart disease based on feed-forward neural network architecture and genetic algorithm. Hybridization has applied to train the neural network using genetic algorithms which is inspired by nature and applied to the field of problem solving, notably function optimization and is proved to be effective in searching large, complex cases. Whenever the doctor submits data, the system identified the data from comparing with the trained data and generates a list of probable diseases that the patient may have. The experimental result shows that the proposed approach is more suitable compare to back propagation which is generally used to train neural network.

A Swine Flu Medical Recommender (SFMR) is developed in [12] with the help of CBR that helps medical officer in detecting swine flu. A case, comprising of several symptoms, each of which are assigned with a weighted value, is submitted as query to the system to compare it with the past cases from its case base using an algorithm for similarity which retrieve the most similar cases. These retrieved cases will be used to produce the probable solution to the problem or may be updated with some modifications.

The paper presented in [13] introduces an approach to implement knowledge within myCBR 3, which focused on vocabulary and similarity measure development system architecture. Vocabulary interprets the range of permissible values for attributes. For numeric values, this is usually the value range (minimum, maximum),

while for symbolic values, this can be a list of values and similarity measures defining the relationship between attribute values. Similarity measures can be calculated using Hamming distance for numeric values or reference tables for symbolic values. The myCBR 3 Workbench supports powerful GUIs for representing the similarity of knowledge. This technique is used in modeling of knowledge model, retrieval of information, and case base management. The idea of global similarity is described after the measurement of local similarity based on available attributes which provides the definition of vocabulary and conceptual idea.

A CBR system for complex medical diagnosis has been discussed in paper [14] in the domain of detection of premenstrual syndrome (PMS) as it falls under both gynecology and psychiatry. To address this issue, the paper has proposed a CBR-based expert system that uses the k-nearest neighbor (KNN) algorithm to search k similar cases based on the Euclidean distance measure. The novelty of the system is in the design of a flexible autoset tolerance (T), which serves as a threshold to extract cases for which similarities are greater than the assigned value of T .

The architecture of the system is shown in Fig. 1. Each part of the system is essentially an individual problem solver. Generally, when the super user submits a case in the form of some symptoms to the system, the system then tries to search for the most similar cases from its existing database called case base using inference engine which processes the knowledge cases to generate a probable solution that helps the super user to identify the problem. The super user only can modify the case base if required.

Fig. 1 Architecture of CEDS

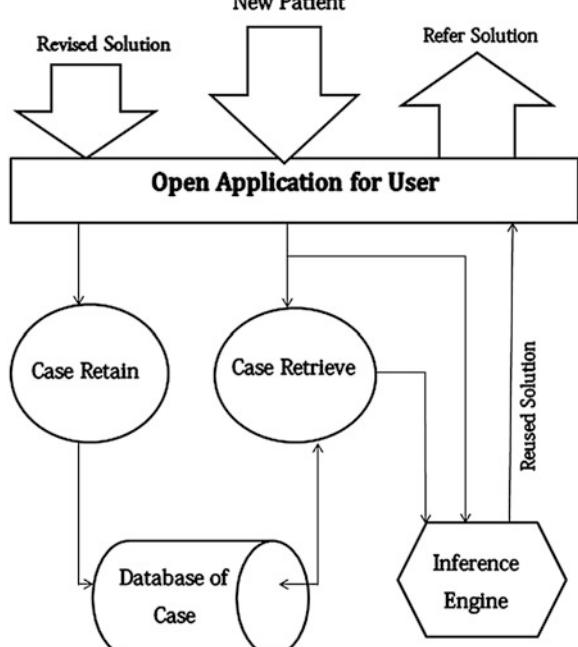


Table 1 Symptoms, weight, and symbols

Symptoms	Weight	Symbol
Vomiting	0.8	VM
Dehydration	0.8	DH
Stomach pain	0.8	SP
Weakness	0.8	WK
Irritability	0.8	IR
Rapid heart rate	0.6	RH
Low blood pressure	0.6	BP
Loss skin elasticity	0.5	SE
Muscle cramps	0.5	MC
Dry muscular membrane	0.6	DM
Thirst	0.3	TH
Weight loss	0.3	WL

2.1 Symptoms, Weight, and Symbol

Each case consists of several symptoms. Each symptom is attached with a weight value which denotes the relative importance of that symptom among the several symptoms exhibited in the disease and intensity of that symptom experienced by the patient expressed in the form of observed value for low, normal, and high. The symptoms are classified as primary (PS) and secondary symptoms (SS) based on list prepared by medical experts as per statistical evidence of the most common symptoms. We set the weight value of each symptom shown in Table 1 on the basis of relative importance values assigned by the medical experts that depend on statistical evidence.

3 Proposed Methodologies

The proposed system CEDS is developed for identifying the cholera disease from the study of several symptom correspondence to that disease using the tools of NetBeans 8.0 and its database. Initially, medical expert submits a well-defined query in the form of symptoms partially collected from the patient. The system works at 2-stage search based on *Cardinality* and *Similarity*, respectively, to analyze the data for results.

In cardinality approach, the system deals with the common primary symptoms with respect to the disease that the patient is suffering from, as in Table 1. On processing the useful data, the CEDS checks whether the given symptoms experienced by the patient match with the maximum number of possible common symptom or not to predict the initial solution as shown in Fig. 2.

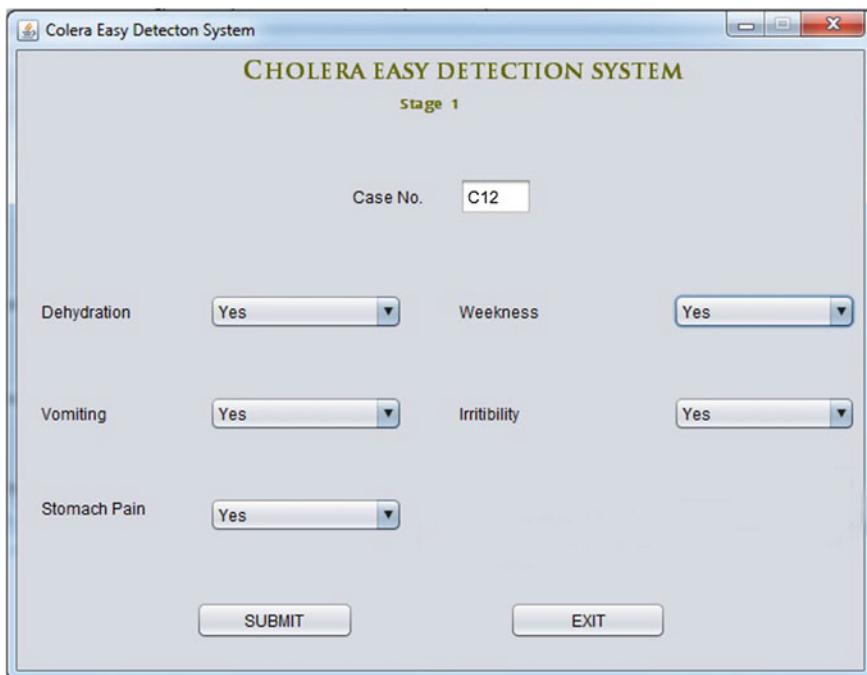


Fig. 2 Symptoms chart with primary symptoms

Table 2 Cardinality approach

Disease	Equations
Stage 1 cholera	$Z^+ = n(z) \geq p$ $Z = \{PS PS \in \text{yes} \text{ and } PS \in C_i\}$ $C_i \rightarrow \text{Set of Symptoms of Case } i$ $p \rightarrow \text{Thresold Value}$

If the desired initial solution is generated, the system will go for next verification to rectify the initial solution (Table 2).

In similarity searching approach, the secondary symptoms are initially fed to the system by the medical officer as per his experience and knowledge in Fig. 3. On processing the data, system will retrieve the similar type of cases from the case database. If C_{new} is the case which is submitted to the system, then the system will return the k most similar cases with respect to threshold value δ_t from its existing case base C as shown in Fig. 4 using k means neighboring approach as shown in Eq. (1). Here, we consider $k = 3$.

The screenshot shows a Windows application window titled "Cholera Easy Detector System". Inside, a title bar says "CHOLERA EASY DETECTION SYSTEM" and "Secondary Stage". Below are eight input fields arranged in two columns. The first column contains: "Age" (23), "Heart Rate" (Rapid), "Thirst" (High), "Weightloss" (20%), "Loss Skin Elasticity" (Yes), and "Water Quality" (Unhygienic). The second column contains: "Sex" (Male), "Muscle Cramps" (Yes), "Dry Mucous Membranes" (90%), "Blood Pressure" (Low), "Food Quality" (Under Cooked), and "Blood Group" (O+). At the bottom are two buttons: "SUBMIT" and "EXIT".

Fig. 3 Symptoms chart with secondary symptoms

CASENO	HEARTRATE	BP	DRYMUSCULERMEMBRENS	WEIGHTLOSS	LOSTSKINELASTICITY	MUSCLECRAMPS	IRRITABILITY	AGE
C24	Rapid	Low	70%	10%	Yes	Yes	No	29
C56	Normal	Low	50%	10%	Yes	Yes	No	42
C63	Rapid	Normal	80%	5%	No	Yes	Yes	47

Fig. 4 Partial case base of similar three cases

$$\text{Sim}(C_i, C_{\text{new}}) = \frac{\sum (O_i^j - O_i^{\text{new}}) \times W_i}{\sum W_i} \quad (1)$$

where O_i^j defined as the observed value O of j th symptom of case i and w is the weight value of that symptom. Now, system will select three most similar cases by Eq. (1) say C_r , C_s , and C_t with respect to C_{new} . Let C' be the most similar cases, then $C' = \{C_p \in \{r,s,t\} | C_p \in C \text{ where } C \text{ set of cases in case base}\}$.

- Calculation of Deviation (DV):

After retrieving the similar cases, system will compare those with the new case C_{new} to check the amount of deviation DV of the individual case from the new as shown.

$$\overline{C_p} = \frac{\sum_{i=1}^{\text{no. of symptoms}} f(C_p; C_{\text{new}}) \times W_i}{\sum W_i} \quad (2)$$

$$\dot{\overline{C_p}} = \frac{\sum_{i=1}^{\text{no. of symptoms}} f^2(C_p; C_{\text{new}}) \times W_i}{\sum W_i} \quad (3)$$

where

$$f(C_p; C_s) = \left| O_p^i - O_{\text{new}}^i \right| \quad (4)$$

where f is the absolute difference of two observed values

$$f^2(C_p; C_{\text{new}}) = \left| O_p^i - O_{\text{new}}^i \right|^2 \quad (5)$$

$$\text{DV}^2(C_p) = \left(\overline{C_p} \right) - X^2 \quad \text{where } X = \overline{C_p} \quad (6)$$

- Best case selection:

The system will return the best case C_{best} by computing the best fitness value by Eq. (7).

$$\text{fitness}(C_p) = \min_{p \in \{r,s,t\}} \{ \text{DV}(C_p) \} \quad (7)$$

If the best case $C_{\text{best}} \geq \tau$, then the patient has a high chance of Cholera. Here, τ is the threshold value.

4 Result and Analysis

$C24$ is the most matched case among the three past similar cases with respect to new case $C12$ which is calculated by the best fitness value of each case by Eq. (7) (Fig. 5).

The deviation of three cases with respect to the new case is represented by three different colors shown in Fig. 6. The vertical axis of the graph shows the deviation

Fig. 5 Report of probable solution with case no.

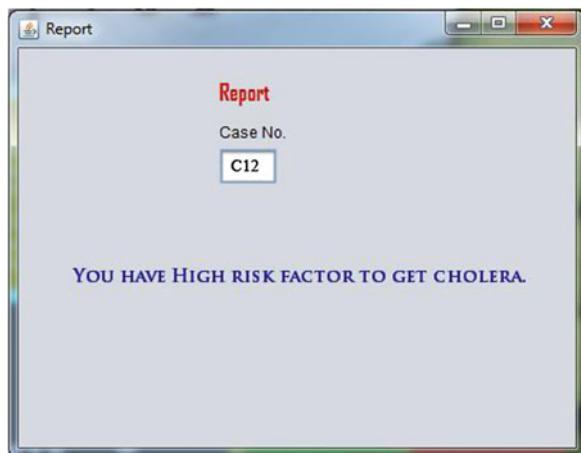
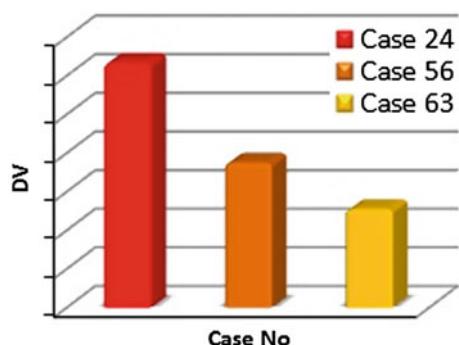


Fig. 6 Comparison of DV of three most similar cases w.r.t new case



(DV) for each case, while the horizontal axis represents the case numbers. The representation of this chart helps the medical expert to take an instant action for diagnosis.

5 Flow Chart

See Fig. 7.

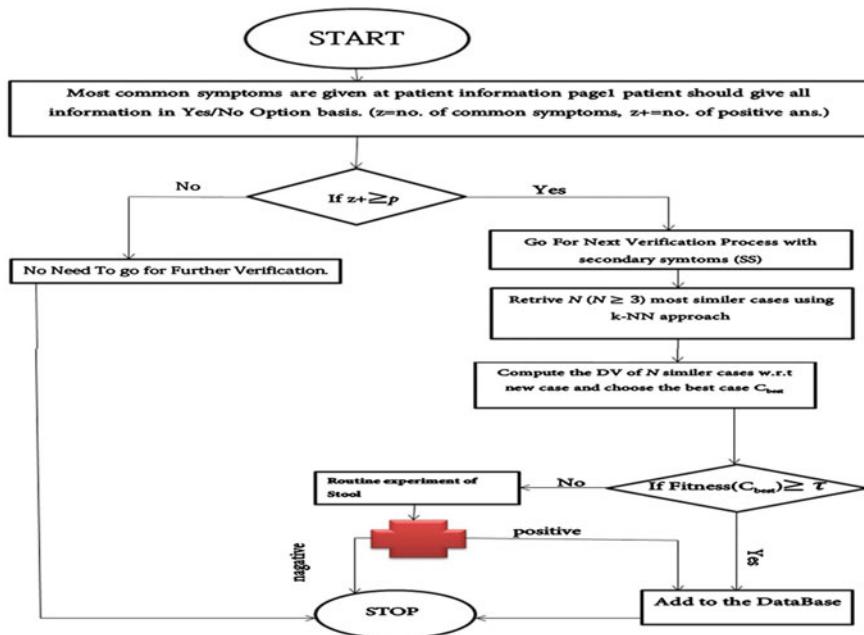


Fig. 7 Flowchart of CEDS

6 Conclusion

In this research, a simple CBR system to detect the cholera diseases has been developed in order to explain how CBR can be helpful in assisting the doctors to take a prominent decision. However, CBR is under the assumption that similar problems have similar solutions, and new problem can be solved by utilizing past similar problems or modifying solutions generated by the system. So a complex CBR case is required before it can match user's condition. In the future, this system can be extended to work in a distributed environment so that it will operate on real-time cases. We also wish to develop our system to work on stronger platform with real-time data.

Acknowledgment The authors express deep sense of gratitude to the JNM Hospital, Kalyani, West Bengal, India, for contributing their valuable advices and to the Dept. of Engineering and Technological studies where the actual work has been carried out.

References

1. Riesbeck, C.K., Schank, R.C.: Inside Case-Based Reasoning. Lawrence Erlbaum Associates Pubs, Hillsdale (1989)
2. Kolodner, J.L.: Case-Based Reasoning. Morgan Kaufmann Publishers Inc., San Francisco (1993)

3. Aamodt, A., Plaza, E.: Case-based reasoning: foundational issues, methodological variations, and system approaches. *Artif. Intell. Commun.* **7**(1), 39–59 (1994)
4. Choy, K.L., Lee, W.B., Lo, V.: Design of case-based intelligent supplier relationship management system—the integration of supplier rating system and product code system. *Expert Syst. Appl.* **25**(1), 87–100 (2003)
5. Miller, P.L.: Selected Topics in Medical Artificial Intelligence: Computer and Medicine. Springer, New York (1998)
6. Cunningham, P., Bonzano, A.: Knowledge engineering issues in developing a case-based reasoning application. *Knowl.-Based Syst.* **12**(7), 371–379 (1999)
7. S-H, Liao: Case-based decision-support system: architecture for simulating military command and control. *Eur. J. Oper. Res.* **123**(3), 558–567 (2000)
8. Dietterich, T., Hild, H., Bakiri, G.: A comparison of ID3 and backpropagation for English text-to-speech mapping. *Mach. Learn.* **18**(1), 51–80 (1995)
9. Deyashi, S., Banerjee, D., Chakraborty, B., Ghosh, D., Debnath, J.: Application of CBR on viral fever detection system (VFDS). IEEE (2011) (978-1-4577-0434-5/11/\$26.00)
10. Kayode, A.A., Afolabi, B.S., Adelusola, K.A.: An integrated knowledge base system architecture for histopathological diagnosis of breast diseases. *Int. J. Inf. Technol. Comput. Sci.* **01**, pp. 74–84. Published Online December 2012 in MECS (<http://www.mecs-press.org/>) doi:[10.5815/ijitcs.2013.01.08](https://doi.org/10.5815/ijitcs.2013.01.08) (2013)
11. Kavitha, K.S., Ramakrishnan, K.V., Singh M.K.: Modeling and design of evolutionary neural network for heart disease detection. *IJCSI Int. J. Comput. Sci. Issues* **7**(5) (2010) (ISSN Online: 1694–0814)
12. Chakraborty, B., Iyenger, S.S., Sood, P., Nabhi, A., Ghosh, D., Debnath, N.: Application of CBR in developing a prototype swine flu medical recommender. In: Special Issue of International Journal of Recent trends in Engineering. Academy Publishers, Finland (2010)
13. Bach, K., Sauer, C., Althoff, K.D., Roth-Berghofer, T.: Knowledge modeling with the open source tool myCBR
14. Chattopadhyay, S., Banerjee, S., Rabhi, F.A., Acharya, U.R.: A case-based reasoning system for complex medical diagnosis. *Expert Syst.* **30**(1), 12–20 (2013)

Efficient Adaptive Routing Algorithm for WSNs with Sink Mobility

Ashwini V. Mannapur, Jayashree D. Mallapur and S.P. Parande

Abstract Some of the application of WSNs such as disaster management and battlefield control demand fast data delivery, adaptable to network dynamism, and recovery from node failure. The proposed work aims overcoming the energy issues, recovers from node failure, and balances the power consumption of nodes while transferring data from source to sink. Optimal route is selected on the basis of minimum distance and minimum transmission energy required. Multi-hop technique is preferred for routing as it consumes less power. Comparing the results with dynamic source routing protocol (DSR), it could be inferred that proposed model is efficient in terms of power consumption; it has better network lifetime and throughput.

Keywords Sensor node · Energy consumed · Adaptive routing

1 Introduction

A wireless sensor networks (WSNs) typically consist of tens to thousands of sensor nodes. These sensor nodes sense and measure physical conditions such as temperature and pressure from the environment, and then, it process and cooperatively pass their collected data to a sink node or base station [1, 2]. These nodes operate on battery, and replacement of it is impractical. The functional life of each individual node varies based upon the demands placed on its battery. Thus, important characteristics in the design of sensor networks are lifetime maximization, robustness, fault

A.V. Mannapur (✉) · J.D. Mallapur · S.P. Parande
Basaveshwar Engineering College, Bagalkot, India
e-mail: ashwini.mannapur@gmail.com

J.D. Mallapur
e-mail: bdmallapur@yahoo.co.in

S.P. Parande
e-mail: somupp_parande@yahoo.com

tolerance, self-configuration, and adaptability. Various routing protocols to save energy consumption and to extend sensor network lifetimes have been studied in [3].

In WSNs, nodes closer to the sink die rapidly due to heavy traffic load for packet transmission which leads to unbalanced power consumption among the sensor nodes; thus, the connectivity within the network may be lost [4]. Although energy is the most critical resource in a WSN, message transfer delay also plays a pivotal role for time critical application such as disaster management.

Most of the existing techniques focus on lifetime maximization of sensor network considering variations such as data aggregation, efficient routing, scheduling, and by introducing sleep mode. However, these techniques compromise on data delivery ratio and latency. In our work, an Efficient Adaptive Routing Algorithm for WSNs with Sink Mobility (EARASM) is proposed. The impulse behind this work is to trim down the energy consumption of sensor node and to allow nodes to dissipate data as and when the data are available using alternative path if the primary path fails due to node failure.

The rest of the paper is organized as follows: Sect. 2 briefs about the related work. Proposed scheme is defined in Sect. 3. Simulation and performance parameters are described in Sect. 4. Section 5 describes the simulation results of our proposed work. Finally, Sect. 6 concludes the paper.

2 Related Work

Extensive research work has been carried out on data routing techniques in WSNs. In [5], authors propose an efficient data forwarding techniques for WSNs. In this paper, they have compared data routing for multiple event occurrence in case of next node with minimum distance (NNMD) and next node with maximum residual energy (NNMRE). The authors in [6] have proposed a dynamic discovers routing method for data transfer between sensor nodes and a sinks in WSN. This method tolerates node failure and small part of network failure. Here, each node records only its neighbor nodes information and adds no extra routing overhead during failure-free periods. They assumed that every node decides its path based only on local information, such as its parent node and neighbor nodes routing information, due to which there is possibility of forming a loop in the routing path.

In [7], authors propose distributed routing algorithm that relies on the local view of a node that relocates the least number of nodes and provides pre-failure intimation of node before failure occurs and ensure that no path between any pair of affected nodes is extended. However, a failure of an actor may cause the network to partition into disjoint blocks. In [8], authors propose a method to maintain network wide energy equivalence and maximize network lifetime so that no critical nodes would become the bottleneck of network lifetime.

In [9], authors have designed technique to maintain the cluster structure in the event of failures caused by energy-drained nodes is proposed. In [10], authors deal with issues that affect the stability and reliability of data transmission in network

layer and present a gradient-based reliable multipath transmission strategy with fault tolerance when the node failure occurs. In [11, 12], authors discuss different methods to deal with node failure and its recovery.

3 Proposed Work

In this section, we present EARASM protocol for WSNs. A source node transmits data packets to sink using one of the routes with low route cost metric. Distance between transmitter and receiver has more effect on energy of the network. Goal of this proposed work is to make network energy efficient and tolerant to node failure.

3.1 Energy Model

Energy model used to calculate the energy consumed by each node to transmit a frame of data to sink. The energy consumed to transmit one bit of data to sink is given by

$$E_b = \frac{-1}{\pi} (1 + \alpha) N_f \sigma^2 \ln(P_b) G_1 M^2 M_1 + \frac{P_r + P_t}{B} \quad (1)$$

And the total number of bits transmitted to the sink by a node in each round is given by

$$S_1 = NF_n PS \quad (2)$$

And the total energy consumed by each node in WSN to transmit data to the sink is given by

$$E_{\text{tot}} = S_1 E_b \quad (3)$$

Using Eq. 3, we can define the energy left in each node after transmission of information is given by

$$\text{Remaining_energy} = \text{Initial_energy} - \text{Energy_consumed} \quad (4)$$

Table 1 describes about the symbol used in equations with their assumed values.

3.2 Route Selection

In a WSN, every source node finds an independent route to the sink and data are transmitted at the rate specified by 802.11 standards from source to sink. Each node

Table 1 Assumed parameter values

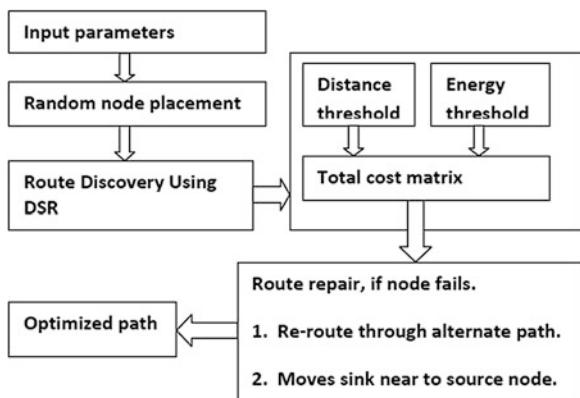
Symbol	Value	Description
α	0.5	Loss factor
N_f	10 dB	Receiver noise figure
$\sigma^2 = \text{No}/2$	-84 dB/Hz	Power density of AWGN channel
P_b	10^{-3}	BER
G_1	40 dB	Gain factor
M_1	30 dB	Gain margin
B	20 kHz	Bandwidth
P_t	98.2 mW	Transmitter power consumption
Pr	112.6 mW	Receiver power consumption
F_n	2	Number of symbols in a frame
P	0.8	Transmit probability of each node
S	2 k bits	Packet size
M	100×100 m	Field area

can communicate to each other over a distance of T_h which is ideally 20 m in real WSN networks. The proposed algorithm is divided into route setup phase and route recovery phase shown in Fig. 1.

For managing WSN, each nodes exchange beacon packets at the beginning. Generally, beacon packets are short. However, as we want to incorporate channel condition in routing, we add payload filed in the beacon packets.

As beacon is transmitted from one node to its nearest node, the receiving node estimates delay, energy consumption, and power loss with respect to each of these nodes and generates route request query (RREQ) packets. Path obtained using dijkstra's shortest path by incorporating energy, power, and delay obtained in the beacon exchange phase.

Fig. 1 Route selection process



The desired condition for best route is to have minimum energy loss and delay with maximum reception power; thus, a route cost matrix is calculated as given below,

$$\text{Cost} = (E_{\text{loss}} * \text{Delay}) / \text{Pr}. \quad (5)$$

where E_{loss} is the energy lost during the transmission of packet, Delay represents time taken for a packet to reach sink from source, and Pr represents receiver circuit power required.

The proposed algorithm obtains the best link from source node from all possible links that has minimum cost in comparison with others. Once route is formed, we transmit round number of data packets from source to sink and measure the parameters along with output parameter throughput.

3.3 Algorithm 1: Route Setup Phase

In this phase, beacon packets are exchanged between sensor nodes and my path table is created.

Note: In our algorithm, we have assumed energy threshold (EN) in the range 80–70 %, and threshold range can be modified to suit real-time applications.

```

1: Begin.
2: Randomly deploy sensor nodes with neighbor list.
3: My path table is created using DSR protocol when
an event occurs.
{
Source node broadcasts beacon (RREQ).
Upon receiving beacon, a node calculates cost metric and
sends back RREP packet.
Cost= (Eloss*Delay)/Pr.
Now based on cost metric of receiver RREP next node is
selected.
If (EN >80% and distance<Th)      //EN energy threshold,
                                         // Th distance threshold.
}
4: Repeat step 3 to construct the routing path from
source node to sink.
5: Store the path details in my path table.
6: After path is found, transmit round number of message
packets from source to sink with beacon (RREQ) as
payload using BPSK model with AWGN.
{
Call algorithm 2
}
7: End.

```

3.4 Algorithm 2: Route Recovery Phase

In this phase, the data are transmitted from source to sink through the route selected and link failure is recovered caused by node failure.

```

1: Begin
2: Every time a node receives a packet || burst of packets, it forwards it to next node.
Node generates RREP and sends back to source node.
3: Based on beacon message received.
< Source node checks: cost metric>
If (EN < 80%) // EN represents energy of node.
{
Call step 3 of algorithm 1 // find new route.
//Update my path table and node transmits complete packet
while new route is found.
//Update status of node to failed (dead).
}
Else If ((EN_source || EN of node) < 70%)
{
Move Sink nearer to Low energy node or dead node.
Update my path table and status of node to dead.}
Else {Continue with existing route.
}
4: Return (parameters).

```

4 Simulation Model

4.1 Simulation

Due to rich function library of MATLAB, it enables easy implement of different types of algorithms for WSN. MATLAB has built-in mathematical functions, graphical functions, etc. MATLAB also allows user to write their own functions and scripts using the built-in functions and data types and data structures. Therefore, MATLAB 12.0 is chosen as simulation environment for implementation of our algorithm for WSN in an area of 100×100 m.

In our paper, we have considered some of the parameters given in Table 2 to illustrate simulation results.

In this paper, BPSK modulation is used to transmit S number of bits from one node to another node. Here, at first, message bits are converted into bipolar signal, and later, the signal is multiplied with a carrier signal. The resultant bits are then

Table 2 Simulation parameters

Parameters	Value
Number of nodes (N)	10–100
Number of messages	10–100 bytes
SNR	–60 to 0
Number of iterations	10–150
Initial energy of node	10 kJ
Transmission range	20 m
Loss factor	0.5

transmitted through Additive white Gaussian noise (AWGN) channel. Ideally in WSN network, SNR sensitivity is about –84 dBm, i.e., two nodes at a distance of 1 m can receive packets even if signal to noise ratio is –84 dB.

4.2 *Simulation Procedure*

We outline the simulation procedure of the proposed system by following steps:

1. Start.
2. Provide network simulation parameters.
3. Generate the WSN for the given area and number of nodes.
4. Apply the proposed Adaptable routing algorithm.
5. Compute the performance of the system.
6. End.

Consider a simple network with 50 nodes where we define some of the parameters of the system such as SNR = –40, message length = 20, iterations = 20 with source id and sink.

4.3 *Performance Parameters*

The performance parameters considered here are as following:

- Average energy consumed: It is defined as the mean of energy consumed for S bit transmission with respect to number of iterations.
- Average BER: It is defined as probability of getting error bits in received packet.
- Average delay: It is defined as the mean of time taken to transmit packet through each link.
- Failure rate: It is defined as number of dead nodes with respect to number of nodes deployed.

5 Results and Discussions

In this section, we discuss various results obtained through simulation. EARASM protocol is compared with existing non-energy aware DSR protocol with static sink.

Figures 2 and 3 depict average energy consumed with respect to increase in number of nodes deployed and number of packets being transmitted, respectively, for proposed scheme and DSR protocol.

In Fig. 2, it can be observed that EARASM protocol consumes 7.6 % average energy compared to 9.4 % for DSR protocol for 10 packets transmitted for 10 times.

In Fig. 3, it can be observed that EARASM protocol consumes 7.9 % average energy compared to 10.38 % for DSR protocol for 50 nodes deployed and simulation is carried out for 10 iterations.

Figure 4 depicts percentage of node failure rate with respect to number of iterations being carried out for 50 sensor nodes deployed in an area of 100×100 m. Results show that the average node failure rate is 10.8 % for EARASM protocol compared to 14.8 % for DSR protocol.

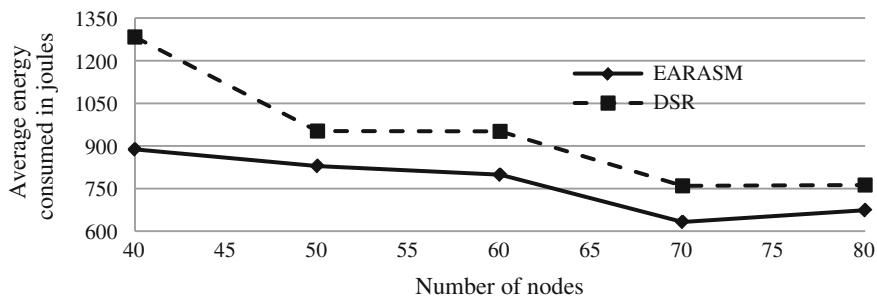


Fig. 2 Number of nodes versus average energy consumed

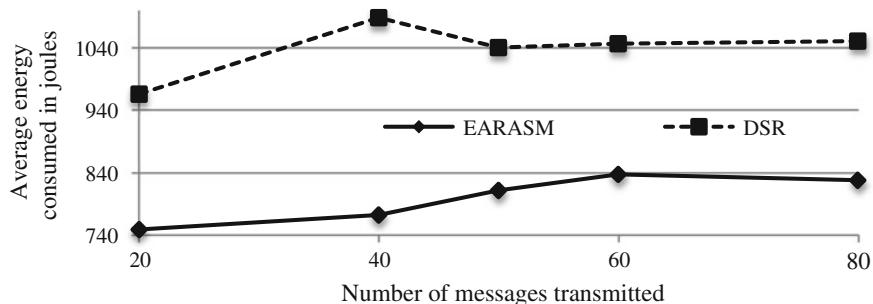


Fig. 3 Number of messages transmitted versus average energy consumed

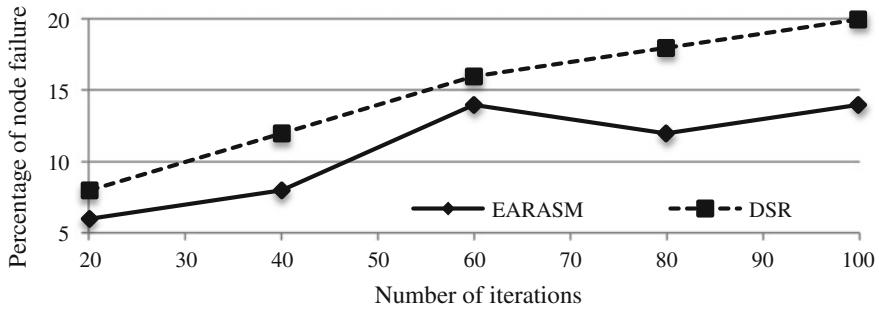


Fig. 4 Number of iterations versus percentage of node failure

Figure 5 depicts average delay obtained for transmitting 10 messages with respect to increasing the number of nodes. It can be observed that average delay is 2.8 s for EARASM protocol compared to 3.7 s for DSR protocol.

Figure 6 presents variation in BER with respect to number of nodes for 10 packets being transmitted for 10 iterations; it can be observed that the EARASM protocol has average of 9 error bits, whereas DSR protocol has average of 14 error bits.

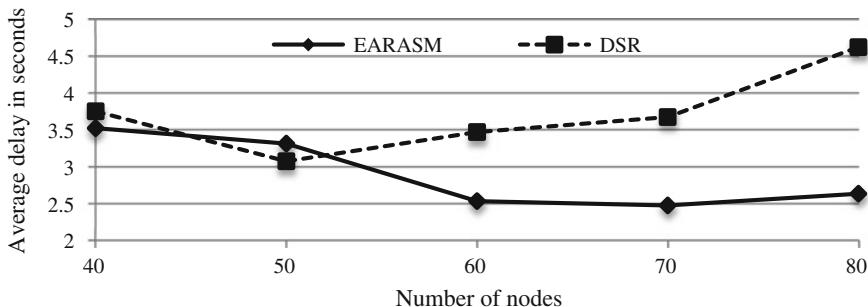


Fig. 5 Number of nodes versus average delay

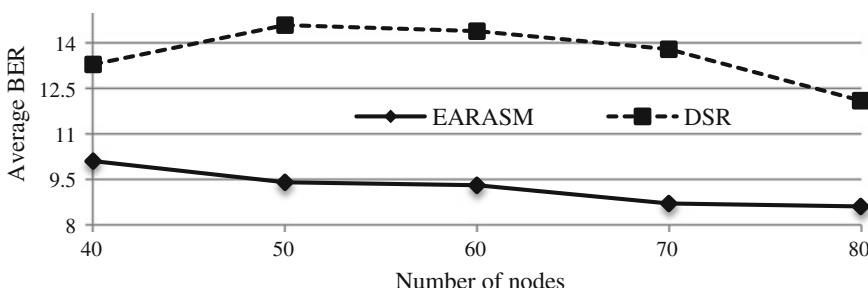


Fig. 6 Number of nodes versus average BER

6 Conclusions and Future Scope

WSN is a modern technology, and it has a very exciting application areas ranging from disaster management to health care and vehicle tracking. Energy efficient network is a vital problem in this field of WSN. Moreover, sensor nodes are prone to failure due to various challenges.

A wireless sensor network links often fail, and the failed sensor nodes cannot be replaced in short time gap, so people must use some energy efficient routing protocol for better result. Proposed scheme is basically a modification of an existing DSR protocol and assures in-time data delivery for critical data.

Proposed scheme consumes nearly 3 % less energy measured against non-energy aware DSR protocol with static sink. It also achieves 4 % less node failure rate, better network lifetime, reduced BER, and transmission delay so that WSNs can be effectively used in real-time applications.

As a future extension of this work, the efficiency of the network can be improved by implementing the proposed algorithm with multiple sink and efficient placement of sinks, which improves lifetime of sensor nodes and ultimately the whole network.

References

1. Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E.: *Wireless Sensor Networks: A Survey*, Elsevier Science, Amsterdam (2002)
2. Maxwell, C.: *A Treatise on Electricity and Magnetism*, vol. 2, pp. 68–73, 3rd edn. Clarendon, Oxford (1892)
3. Garg, P., Rani, R.: A survey on wireless sensor networks routing algorithms. *IJITKMSpecial Issue, ICFTEM*, 38–42 (2014). ISSN 0973-4414
4. Akan, O.B., Akyildiz, I.F.: Event -to-sink reliable transport in wireless sensor networks. *IEEE/ACM Trans. Netw.* **13**(5), 1003–1016 (2005)
5. Rehena, Z., Roy, S., Mukherjee, N.: Efficient data forwarding techniques in wireless sensor networks. In: *3rd IEEE International Advance Computing Conference (IACC)*, pp. 449–457 (2013)
6. Nanda, A., Rath, A.K., Rout, S.K.: Node sensing and dynamic discovering routes for wireless sensor networks. *Int. J. Comput. Sci. Inform. Secur.* **7**(3), 121–131 (2010)
7. Sureka, M., Jany Shabu, S.L.: Pre-recover from a node failure in Ad hoc network using reactive protocols. *Int. J. Eng. Trends Technol.* **5**(8), 262–266 (2014)
8. Ding, W., Iyengar, S.S., Kannan, R., Rummler, W.: Energy equivalence routing in wireless sensor networks. *Microprocess. Microsyst.* **28**, 467, 475 (2004)
9. Akbari, A., Dana, A., Khademzadeh, A., Beikmahdavi, N.: Fault detection and recovery in wireless sensor network using clustering. *Int. J. Wirel. Mobile Netw.* **3**(1), 130–138 (2011)
10. Yang, J., Ye, Z., Ouyang, Z.: A strategy for fault recovery of wireless sensor network based on v-SVC. *J. Inform. Comput. Sci.* **10**(6), 1693–1704 (2013)
11. Belghachi, M., Feham, M.: QoS routing scheme and route repair in WSN. *Int. J. Adv. Comput. Sci. Appl.* **3**(12), 81–86 (2012)
12. Liu, X., Zhao, H., Yang, X., Li, X., Wang, N.: Trailing mobile sinks: a proactive data reporting protocol for wireless sensor networks. In: *National Science Foundation (grants CNS-0709329 and CNS-0923238) and OCAST, 2010*

Detection of Degree of Sickness of Affected Eye Using Fuzzy Logic and Histogram Analysis

Manisha Barman, J. Paul Choudhury and S. Biswas

Abstract Cataract is a clouding of the lens inside the eye which leads to a decrease in the vision. It is the most common cause of blindness and is conventionally treated with surgery. The main objective of this paper is to develop a system which helps to detect cataract in affected eyes. The system consists of selection of area of interest from an image of an image and extraction of certain features of images. The proposed work consists of selection of the area of interest from an eye image and formation of membership values and their mean values from the textual properties of the image with an objective to differentiate between the normal and affected eye for the purpose of detection of degree of sickness of the eye. The said work is confirmed from the result as obtained from the irregular frequency distribution of the input image and the value of irregular deviation of the frequency of the image. A set of images consisting of normal and affected eyes has been collected for carrying out this work.

Keywords Cataract eye · Membership function · Irregular distribution · Mean value · Irregular deviation

1 Introduction

A cataract is a clouding of the lens inside the eye which leads to a decrease in vision. It is the most common cause of blindness and is conventionally treated with surgery. Age-related cataracts are responsible for 51 % of world blindness, about

M. Barman (✉) · J.P. Choudhury
Department of Information Technology,
Kalyani Government Engineering College, Kalyani, Nadia, India
e-mail: memanisha5@gmail.com

J.P. Choudhury
e-mail: jnpsc193@yahoo.com

S. Biswas
Department of Engineering and Technological Studies,
University of Kalyani, Kalyani, Nadia, India
e-mail: biswas.su@gmail.com

20 million people. Globally, cataracts cause moderate-to-severe disability in 53.8 million (2004), 52.2 million of whom are in low- and middle-income countries. In many countries, surgical services are inadequate, and cataracts remain the leading cause of blindness. Even where surgical services are available, low vision associated with cataracts may still be prevalent as a result of long waits for, and barriers to, surgery—such as cost, lack of information, and transportation problems. This problem is largely seen in the rural section of India so we decided to work on this kind of abnormality.

The colored part of the eye controls the amount of light that enters into the eye. It is the most visible part of the eye. It is called iris. The iris is divided into two zones: pupillary zone, the inner part of the iris that forms the pupil's boundary, and ciliary zone, the remaining part of the iris that extends into the ciliary body. The color of the iris is determined by a dark pigment called melanin.

Danciu et al. [1] have used a feature extraction approach inspired by the scale-invariant feature transform (SIFT) algorithm. They have used color gradient based on principal component analysis (PCA) to discover the correct key points position, own color corner detection algorithm, for getting result they used images of human eye iris affected by melanoma and also on images representing color textures of healthy iris.

Senthilkumaran and Rajesh [2] have used a survey based on the theory of edge detection for image segmentation using soft computing approaches such as fuzzy logic, genetic algorithm, and neural network.

Adeli and Neshat [3] have used a fuzzy expert system for the diagnosis of heart disease. The authors have used several variables, viz. as chest pain type, blood pressure, cholesterol, resting blood sugar, resting maximum heart rate, sex, electrocardiography (ECG), exercise, old peak (ST depression induced by exercise relative to rest), thallium scan, and age as inputs. The status of the patients as healthy or sick has been used as output. Four types of sickness have been used as output. These are Sick s1, Sick s2, Sick s3, and Sick s4.

Tobias and Seara [4] have used an approach to threshold the histogram according to the similarity between gray levels. Such a similarity is assessed through a fuzzy measure. In this way, they have overcome the local minima that affect most of the conventional methods. The experimental results have demonstrated the effectiveness of the proposed approach for both bimodal and multimodal histograms.

Ercal et al. [5] have proposed a simple and effective method to find the borders of tumors as an initial step toward the diagnosis of skin tumors from color images. The method has made use of an adaptive color metric from the red, green, and blue (RGB) planes that contain information to discriminate the tumor from the background. Using this suitable coordinate transformation, the image has been segmented. The tumor portion has been extracted from the segmented image, and borders have been drawn.

Xu et al. [6] have developed a three-step segmentation method using the properties of skin cancer images. The steps of this method are as follows: (1) preprocessing: A color image has been first converted into an intensity image in such a way that the intensity of a pixel can illustrate the color distance of that pixel with the color of the background. The color of the median of a set of pixels in small window of the images has been taken as the color of the background. (2) Initial segmentation: A threshold value has been determined from the average intensity of high-gradient pixels in the obtained intensity image. To find approximate lesion boundaries, this threshold value has been used. (3) Region refinement: Using edge information of the image, a region boundary has been refined. This has involved in initializing a closed elastic curve at the approximate boundary, and shrinking and expanding it to fit to the edges in its neighborhood.

A fuzzy rule base system has been designed for the detection of heart disease [7, 8]. The developed system has seven inputs. These are chest pain type, resting blood pressure in mm (Trestbps), serum cholesterol in mg (Chol), numbers of Years as a smoker (years), fasting of blood sugar (fbs), maximum heart rate achieved (thalach), and resting blood rate (trestbpd). The angiographic disease status of heart of patients has been recorded as output. It is to state that diagnosis of heart disease by angiographic disease status is assigned by a number between 0 and 1, and that number indicates whether the heart attack is mild or massive.

The work related to detect the sickness of eye (cataract) using image processing system cannot be done earlier. That is the reason to carry out the research work in this line. Here, it is proposed to take images of eyes from normal and sick persons. Thereafter, region of interest from the image of eye has to be taken. Membership functions under fuzzy logic have to be applied on the pixel value of image. After selecting the suitable membership function, column-wise mean value of membership has to be calculated. From the mean value of the degree of cataract can be detected. The said result has to be cross-checked by applying histogram analysis of the pixel values and making the irregular deviation of the frequency distribution value.

2 Methodology

In proposed system, two different approaches have been used. One is using membership function under fuzzy logic. The proposed system has produced membership values and calculated their mean value. On the other hand, histogram analysis has been used to find out their frequency distribution. The derived irregular distribution has used to distinguish between two sets of images.

The flow diagrams of two approaches are given in Figs. 1 and 2.

Fig. 1 Flow diagram of approach I

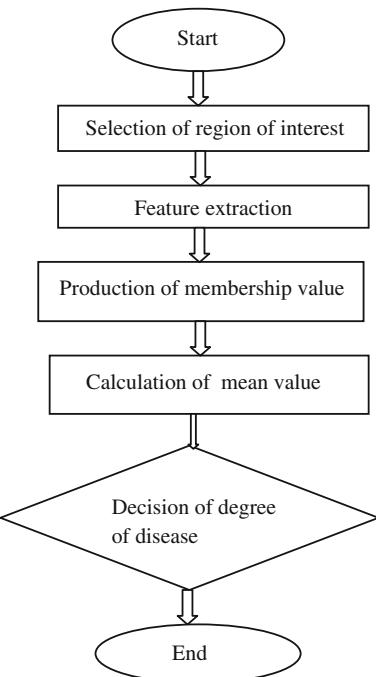
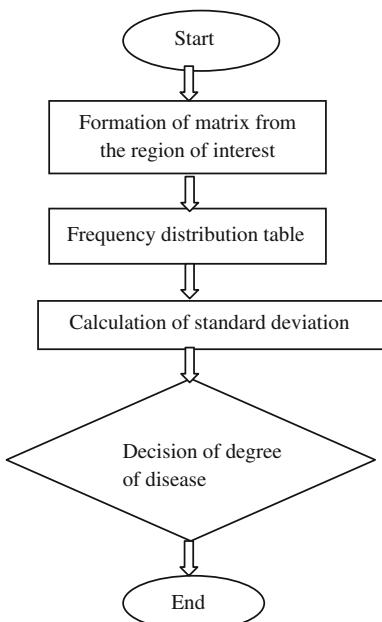


Fig. 2 Flow diagram of approach II



2.1 Fuzzy Membership Function

2.1.1 Z-shaped Base Membership Function

The Z-shaped base membership function is a spline-based function of μ_A of vector x and depends on two scalar parameters p and q that locate the extremes of the sloped portion of the curve. The Z-shaped base membership function μ_A of vector x has been represented by

$$\mu_A(x : p, q) = \begin{cases} 1 & \text{if } x \leq p \\ 1 - 2\left(\frac{x-p}{q-p}\right)^2 & \text{if } p \leq x \leq p + q/2 \\ 2\left(\frac{x-q}{q-p}\right)^2 & \text{if } (p+q)/2 \leq x \leq q \\ 0 & \text{if } x \geq q \end{cases}$$

2.1.2 Sigmoid-based Membership Function

The sigmoid function, $\mu_A(x, [p, r])$, as given in the following equation by $\mu_A(x, p, r)$, is a mapping on a vector x and depends on two parameters p and r . Depending on the sign of the parameter p , the sigmoidal membership function is inherently open to the right or to the left and thus is appropriate for representing concepts such as “very large” or “very negative.” More conventional-looking membership functions can be built by taking either the product or difference of two different sigmoidal membership functions

$$\mu_A(x : p, r) = \frac{1}{1 + e^{-p(x-r)}}$$

2.1.3 Gaussian Combination Membership Function

$$\mu_A(x; \sigma, c) = e^{\frac{-(x-c)^2}{2\sigma^2}}$$

The function gauss2mf is a combination of two parameters sig and c . The first function, specified by sig1 and $c1$, determines the shape of the leftmost curve. The second function specified by sig2 and $c2$ determines the shape of the rightmost curve. Whenever $c1 < c2$, the gauss2mf function reaches a maximum value of 1. Otherwise, the maximum value is less than one.

2.2 Irregular Distribution

From the histogram of image, irregular deviation has been calculated using following steps.

1. If h_i be the amplitude of histogram at index i , then the mean value M_H is given as follows:

$$M_H = \frac{1}{256} \sum_{i=0}^{255} h_i \quad (1)$$

2. Absolute of histogram derivations from this mean value H_{Di} is given as follows:

$$H_{Di} = |h_i - M_H| \quad (2)$$

3. Irregular Deviation I_D is given as follows:

$$I_D = \sum_{i=0}^{255} H_{Di} \quad (3)$$

From this irregular deviation whether the taken two samples are both of same kind or not have been calculated.

3 Implementation

Step 1: Conversion of RGB image to gray image

Eye images have been collected from the CASIA [9]. Snapshot of the selected sample input images in the original in RGB form has been converted into gray form. The corresponding pixel value has been used in step 2.

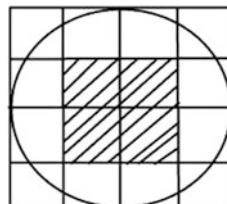


Fig. 3 Selection of region of interest

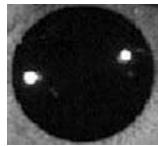


Fig. 4 Sample taken



Fig. 5 Region of interest

Step 2: Selection of Region of Interest

A lot of samples have been taken to find out some deformity in these samples.

From the selected image, the region of interest has been taken. Original sample has been broken into four coordinates, and further, each coordinate also has been broken into four coordinates; thereafter, region of interest has been taken as furnished in Fig. 3.

By using above process, the redundant part from the sample image has been eliminated (Figs. 4 and 5).

Step 3: Calculation of Fuzzy Membership value

Different membership functions (Zmf, Smf, gauss2mf, etc.) have been applied on the pixel values of the region of interest of selected image. It has been observed that Z membership function has produced best result on basis of minimum difference of the fuzzy original value of pixel value and the membership value as obtained.

Step 4: Calculation of Mean Value

The mean value of the membership value column wise has been calculated.

Step 5: Calculation of irregular deviation using histogram analysis generating frequency distribution table

Histogram analysis has been carried out on the region of interest of the image. From histogram analysis, frequency distribution table has been formed.

Step 6: Irregular deviation of the frequency has been calculated from the frequency distribution table.

- Step 7: The instructions as narrated in step 1 to step 6 have been executed for two class images (normal eye and eye with cataract). From the mean as calculated in step 4 and the irregular deviation as calculated in step 6, the degree of sickness (cataract affected) of eye can be detected.
- Step 8: The analysis as made from the value as obtained in step 4 can be confirmed from the analysis as made from the values as obtained in step 6.

4 Result

It has been observed that mean value of normal eye lies between 0.919 and 0.992 and that of affected eye lies between 0.525 and 0.766. The irregular deviation of normal eye lies between 1,031.5 and 1,192 and that of affected eye lies between 344 and 875.35. The computation of mean value for normal eye and cataract-affected eye has been furnished in Tables 1 and 2, respectively.

The computation of irregular deviation for normal eye and cataract-affected eye has been furnished in Tables 3 and 4, respectively.

Table 1 Mean value for normal eye

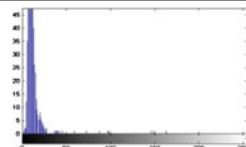
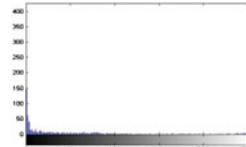
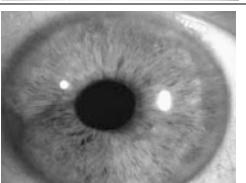
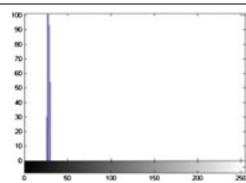
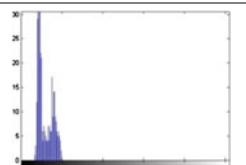
Sample	Histogram	Mean value
		0.9917
		0.9596
		0.9761
		0.9191

Table 2 Mean value for defective eye

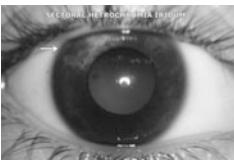
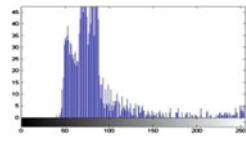
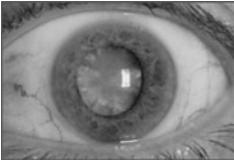
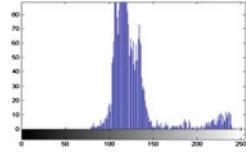
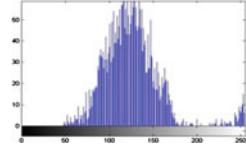
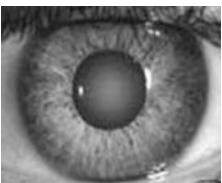
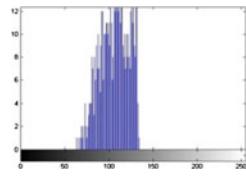
Sample	Histogram	Mean value
		0.7662
		0.5262
		0.5251
		0.6424

Table 3 Irregular deviation for two normal eyes

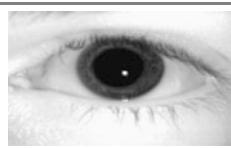
Sample 1	Sample 2	Irregular deviation
		1,192
		1,031.533

Table 4 Irregular deviation for a normal and a defective eye

Sample 1	Sample 2	Irregular deviation
		875.3514
		344

5 Conclusion

In this proposed technique, sample images are taken from normal and sick persons. The images may be distorted due to the heterogeneous communication media and for using different computer platforms. Therefore, it is necessary to rectify the image as good as possible so that the noise can be eliminated, the image will be of good quality, and the available pixel value will be nearest to its original value. That is the reason for using fuzzy logic.

After getting region of interest of images, it is necessary to remove noise from the images based on the pixel value. Applying proper fuzzy membership value, it is desirable that the minimum changes should make the selected region of interest more suitable to work. Three types of fuzzy membership functions have been applied for proper selection of membership functions.

Analyzing various images, it has been observed that if the mean value is 0.996 (approx), the condition of eye is all right; if it is 0.96 (approx), existing condition of eye is all right, but in course of time, it may lead to cataract. Similarly, if the mean value is 0.75 (approx), the eye has been attacked by cataract, but in a mild condition, whereas if it is 0.64, the cataract condition is massive.

Similarly, the value of irregular deviation is 1,032, although the condition is all right at present, but in course of time, it may affect cataract. Similarly, if the value of irregular deviation is 874, the eye has been attacked by cataract, but the degree of sickness is mild. By using the proposed methodology, the analysis as obtained from one method (calculation of mean) can be cross-checked by the analysis as obtained from another method (calculation of irregular deviation). The condition of eye of new person can be detected by using this proposed methodology.

To take betterment of the result, some other soft computing models such as particle swarm optimization algorithm, ant colony optimization algorithm, tabu search method, and firefly algorithm can be attempted for more accurate performance.

References

1. Danciu, G., Banu, S.M., Ivanovici, M.: Scale and rotation-invariant feature extraction for color images of iris melanoma. In: IEEE 2012, pp. 8–12. ISBN 978-1-4673-1653
2. Senthilkumaran, N., Rajesh, R.: Edge detection techniques for image segmentation —a survey of soft computing approaches. *Pattern Recogn.* **37**(6), 1189–1200 (2004)
3. Leeuwen, J.V.: Computer science today. Recent trends and developments. In: Lecture Notes in Computer Science, vol. 1000. Springer, Berlin, Heidelberg (1995)
4. Tobias, O.J., Seara, R.: Image segmentation by histogram thresholding using fuzzy sets. *IEEE Trans. Image Process.* **11**(12), 1457–1465 (2002)
5. Ercal, F., Moganti, M., Stoecker, W.V., Moss, R.H.: Detection of skin tumor boundaries in color images. *IEEE Trans. Med. Imaging* **12**(3), 1–8 (1993)
6. Xu, L., Jackowski, M., Goshtasby, A., Roseman, D., Bines, S., Yu, C., Dhawan, A., Huntley, A.: Segmentation of skin cancer images in image and vision computing. *IEEE Trans. Med. Imaging* **17**, 65–74 (1999)
7. Barman, M., Paul Choudhury, J.: A fuzzy rule base system for the diagnosis of heart disease. *Int. J. Comput. Appl.* (0975 – 8887) **57**(7), 46–53 (2012)
8. Barman, M., Paul Choudhury, J.: A framework for selection of membership function using fuzzy rule base system for the diagnosis of heart disease. *Int. J. Inf. Technol. Comput. Sci.* **11**, 62–70 (2013). doi:[10.5815/ijitcs.2013.11.07](https://doi.org/10.5815/ijitcs.2013.11.07). Published online Oct 2013 in MECS (<http://www.mecs-press.org/>)
9. CASIA-irisv3, nst. of Automation, Chinese Academy of Sciences
10. Cao, W., Che, R., Ye, D.: An illumination-independent edge detection and fuzzy enhancement algorithm based on wavelet transform for non-uniform weak illumination images. *Pattern Recogn. Lett.* **29**(3), 192–199 (2008)
11. Canny, J.: A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **8**(6), 679–698 (1986)
12. Daugman, J., Downing, C.: Epigenetic randomness, complexity, and singularity of human iris patterns. *Proc. R. Soc. Lond. B Biol. Sci.* **268**, 1737–1740 (2001)

Element Spacing Optimization of Low Sidelobe Concentric Hexagonal Antenna Arrays Using MOEA/D

Sudipta Das, Durbadal Mandal, Rajib Kar and Sakti Prasad Ghoshal

Abstract MOEA/D is a popular decomposition based on multiobjective evolutionary algorithm. This work considers a triobjective concentric hexagonal array geometry optimization problem where the three objectives are to obtain low sidelobe radiation pattern with high directivity and with as low number of active elements as possible. For simplicity, the optimization problem is modified from min-max-min search to min-min-min search using the duality principle. An unconstrained search utilizing MOEA/D based on Tchebyshoff decomposition method is conducted to obtain the optimal front. Results reflect that the obtained optimal geometries perform much better than several regularly spaced larger concentric regular hexagonal antenna arrays.

Keywords Multiobjective evolutionary algorithm · MOEA/D · Low sidelobe synthesis · Element spacing optimization · Concentric hexagonal antenna arrays

1 Introduction

Real-world problems are complex, and most of the time, they are related to other problems. Need to find a solution that simultaneously solves all such problems to the acceptable levels becomes evident. Problems can be related in two ways. In the

S. Das (✉) · D. Mandal · R. Kar

Department of Electronics and Communication Engineering, National Institute

of Technology Durgapur, Durgapur 713209, India

e-mail: sudipta.sit59@gmail.com

D. Mandal

e-mail: durbadal.bittu@gmail.com

R. Kar

e-mail: rajibkarece@gmail.com

S.P. Ghoshal

Department of Electrical Engineering, National Institute of Technology Durgapur,
Durgapur 713209, India

e-mail: spghoshalnitdgp@gmail.com

first type of co-relation, solving one problem automatically solves the others, and this is the most favorable scenario for a user. A single-objective optimization will suffice in this case. The other type of correlation is improving quality solution for one problem causes worsening of the other. The smallest set of solutions that best manages this relation is named as Pareto optimal set of solutions (PS). None of the solutions in a PS can be found to provide better response for all problems than any other solution in the same PS. Map of PS in the objective domain is called the Pareto optimal front (PF). Consequently, two basic goals of to obtain solutions in PF are as follows: to obtain the set of best possible combinations of outcomes and to sample the PF as most evenly as possible.

According to a recent survey on multiobjective evolutionary algorithm (MOEA) [1], research for designing reliable structures has been conducted in six ways: (i) MOEA based on decomposition [2], (ii) MOEA based on preference [3–6], (iii) indicator-based MOEAs [7–9] (iv) hybrid MOEAs [2, 10–12], (v) memetic MOEAs [13–14], and (vi) MOEA based on coevolution [15].

The objective of the present work was to obtain an optimal front of three parameters of concentric regular hexagonal antenna arrays (CRHAA) [17–20], named as sidelobe level (*SLL*), directivity (*D*) [16], and number of elements (*N*). This work models the problem as unconstrained optimization problem. This paper utilizes MOEA/D for obtaining this front.

The flow of the present work, elaborated briefly in the next sections, is given as follows: Sect. 2 briefly discusses the problem definition, Sect. 3 briefly discusses the steps of MOEA/D, Sect. 4 records the results and related discussions, and Sect. 5 concludes on the present work.

2 Problem Definition

2.1 The Array Factor

Peak interference–directivity profile of antenna arrays is important for antenna engineers. This work focuses on obtaining this minimum peak sidelobe versus maximum directivity front for hexagonal arrays. A concentric regular hexagonal array is shown in Fig. 1.

Array factor of CRHAA can be given as below [20]:

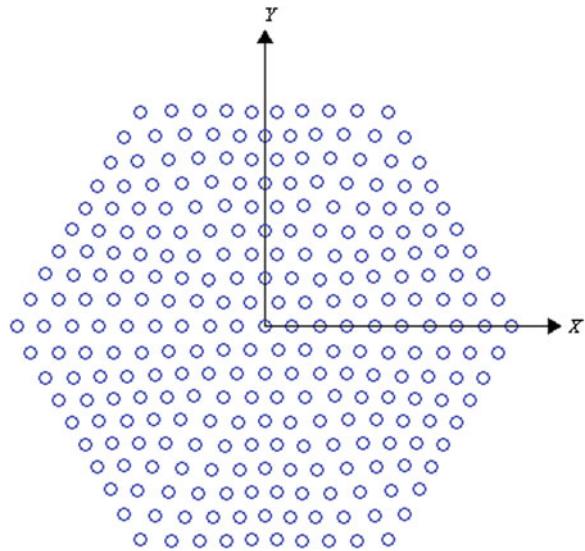
$$AF = 1 + \sum_{m=1}^M \sum_{n=1}^{N_m} \sum_{q=1}^6 e^{jkr_{mnq}} \sin \theta \cos(\phi - \phi_{mnq}) \quad (1)$$

where

N_m is the number of elements in one side of a triangular sector on m th ring; Thus, $N = 1 + 6 \sum_m N_m$

k is the wave number;

Fig. 1 Geometry of a uniform CRHAA



$\{\theta, \phi\}$ is the angular coordinate of all the points in the visible region of the array, $\{\theta \in (0, \pi), \phi \in (0, 2\pi)\}$;

$\{r_{mnq}, \phi_{mnq}\}$ are the coordinates of l th element of n th side of m th ring in one angular sector of the array geometry. Given radius a_m and interelement gap for a ring d_m for m th ring, r_{mnq} , ϕ_{mnq} and N_m can be calculated as follows:

$$N_m = \left\lfloor \frac{a_m}{d_m} \right\rfloor$$

$$r_{mnq} = \sqrt{a_m^2 + (n-1)^2 d_m^2 - a_m(n-1)d_m}, \quad n = 1, 2, 3 \dots N_m \quad (2)$$

$$\phi_{mnq} = \cos^{-1} \left(\frac{r_{mnq}^2 + r_{mnq-1}^2 - d_m^2}{2r_{mnq}r_{mnq-1}} \right)^{\left(\frac{1}{2}\right)} + \frac{(q-1)\pi}{3}$$

For example, a 9-ring CRHAA geometry with $a_m = \frac{m\lambda}{2}$, $d_m = \frac{\lambda}{2}$ has $N = 271$. Graphical illustration of array factor is given in $u-v$ space, where $u = \sin \theta \cos \phi$ and $v = \sin \theta \sin \phi$.

2.2 The Objective Function

Goal of this work was to obtain an optimal design of concentric hexagonal array that can provide possible array designs with the lowest sidelobe pencil beam pattern

and the highest directivity with the smallest number of elements on the aperture. The objective function vector or cost function vector (\overrightarrow{CF}) is given as follows:

$$\overrightarrow{CF} = \{\text{SLL}, D', N\} \quad (3)$$

In (3), D' refers to the negative of peak directivity D . D is usually positive quantity, but for the sake of simplicity, a minus sign is associated with it so as to change the min–max–min objective vector to min–min–min objective vector.

Since interrelation between the parameters of \overrightarrow{CF} is not yet mathematically modeled, authors opted to carry out unconstrained search.

3 Multiobjective Evolutionary Algorithm Based on Decomposition (MOEA/D)

MOEA/D [2] is proposed to handle a multiple objective problem (MOP) by decomposing MOP in several single-objective subproblems. Among several decomposition methods available [21], this work utilizes weighted Tchebysheff decomposition approach. For an \mathfrak{M} -objective MOP, if a L is a vector of \mathfrak{M} weights to combine MOP to form a single-objective subproblem, Tchebysheff approach returns

$$g^{te}(x|L, z^*) = \max_{1 \leq i \leq \mathfrak{M}} \left\{ L_i \left| \frac{y_i - z_i^*}{y_i^{\text{nad}} - z_i^*} \right| \right\} \quad (4)$$

where $y_i = f_i(x|x \in \Omega)$, is the objective value of i th objective, y_i^{nad} is the nadir vector in PS , $z^* = (z_1^*, z_2^*, \dots, z_{\mathfrak{M}}^*)$, $z_i^* = \max\{f_i(x)|x \in \Omega\}$, $1 \leq i \leq \mathfrak{M}$, and $\sum_{i=1}^{\mathfrak{M}} L_i = 1$. The steps of real-coded genetic algorithm based on MOEA/D are given below:

1. Initialize a population of \aleph solutions ($\vec{x}^j, 1 \leq j \leq \aleph$) in d dimensional solution space ($x_j^i, 1 \leq j \leq d$) within search limits and \aleph weights in \mathfrak{M} -dimensional weight space; constraints; An empty External Population (EP), Internal parameters of Genetic Algorithm and Stopping criteria.
2. Compute neighbor matrix $B(j), 1 \leq j \leq \aleph$ such that for each j , it contains the indices of T closest weight neighbors.
3. Until stopping criteria is satisfied, run the following steps:
4. For each $j \in [1, \aleph]$
 - (a) Apply genetic operators to create offspring solution \vec{v}^j and repair using problem-specific method.
 - (b) Update z^* and z^{nad} .

- (c) Replace parent solution if $g^{te}(y_j|L_j, z^*) \leq g^{te}(x^j|L_j, z^*)$.
- (d) Remove all the solutions from EP dominated by \vec{v}^j . If \vec{v}^j is not dominated by any solutions remaining in EP , include \vec{v}^j in EP .

For this work, real-coded genetic algorithm based on MOEA/D is adapted for obtaining optimized front. Simulated Binary Crossover operator with Crossover parameter 2, and Polynomial Mutation operator with mutation parameter 2 are selected. For this work, M is 3, d is selected as 12 (i.e., 6-ring array), population size N is selected 50, and stopping criteria is selected as 200 generations. z^{had} is initialized as $\{-17.54, 0, 200\}$.

4 Simulation Results and Discussion

All simulation is conducted using MATLAB 7.5 software on CoreTM2 Duo processor, 2.99 GHz with 1-GB RAM.

Table 1 records the three optimized array geometries (a_m, d_m) with their respective SLL , D and N values. Both a_m and d_m are expressed in λ , and both SLL and D are expressed in dB. Figure 2 portrays the obtained Pareto front in this work. Figure 3 portrays the radiation pattern of the first set of the tabulated CRHAA geometries.

The first set performs better than others in terms of SLL value, the second array performs better than others in terms of D value, and the third set performs better than others in terms of N value. All these sets can be seen to be non-dominated w.r.t. one another. A 9-ring CRHAA with $a_m = \frac{m\lambda}{2}$, $d_m = \frac{\lambda}{2}$ has 271 elements with SLL value of -18.25 dB and D value of 28.57 dB. To this end, all the evolved arrays could be seen to perform better at least in terms of SLL and N parameters than the aforesaid 9-ring CRHAA in almost every aspect.

Table 1 Geometries and pattern parameters of CRHAA designs evolved with MOEA/D

Set no.	Array geometries						SLL	D	N	
	m	1	2	3	4	5	6			
1	a_m	1.00	1.50	2.08	2.89	3.82	4.61	-22.60	28.44	145
	d_m	0.50	0.58	0.51	0.50	0.57	0.78			
2	a_m	0.91	1.61	2.37	3.21	4.21	5.01	-22.22	29.05	151
	d_m	0.50	0.50	0.50	0.57	0.63	0.73			
3	a_m	0.89	1.40	1.92	2.48	3.36	4.14	-20.49	27.67	127
	d_m	0.50	0.50	0.53	0.66	0.50	0.64			

Fig. 2 Pareto front obtained using MOEA/D

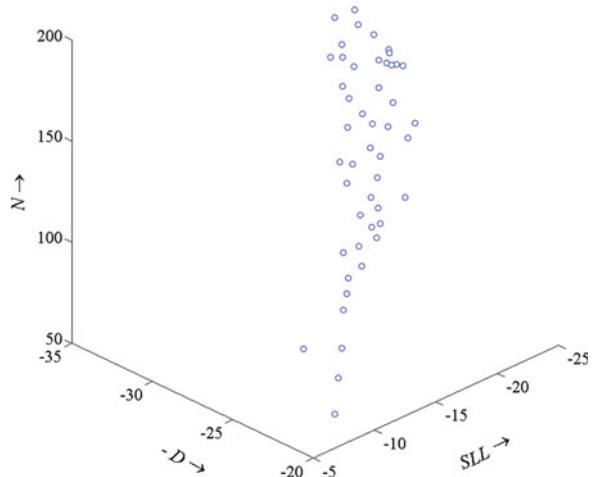
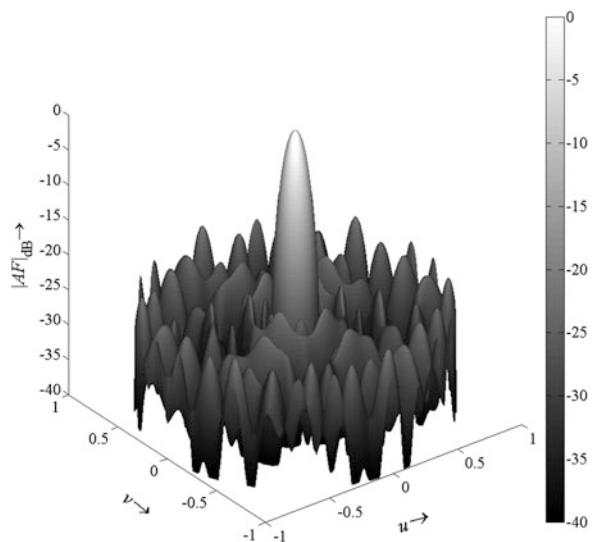


Fig. 3 Array factor of CRHAA geometry (set 1 of Table 1) in $u-v$ space



5 Conclusions

From the results, it can be seen that the obtained CRHAA geometries are able to produce similar radiation profiles as a nine-ring CRHAA with element gap of $\frac{\lambda}{2}$. Evolved array geometries are better than aforesaid array in almost every aspect. For simplicity, this work considers only ideal elements, but if the practical elements were considered, decisions would have been made based on the obtained pattern parameters, and thus, the performance of MOEA/D would not hamper. MOEA/D thus can be successfully applied to search in such difficult trade-off objective space.

6 Acknowledgments

This work is supported by SERB, Department of Science and Technology, Government of India (project no. SB/EMEQ-319/2013).

References

1. Zhou, A., Qu, B.-Y., Li, H., Zhao, S.-Z., Suganthan, P.N., Zhang, Q.: Multiobjective evolutionary algorithms: a survey of the state of the art. *Swarm Evol. Comput.* **1**, 32–49 (2011)
2. Zhang, Q., Li, H.: MOEA/D, a multiobjective evolutionary algorithm based on decomposition. *IEEE Trans. Evol. Comput.* **11**(6), 712–731 (2007)
3. Fonseca, C.M., Fleming, P.J.: Genetic algorithms for multiobjective optimization: formulation, discussion and generalization. In: 5th International Conference Genetic Algorithms, pp. 416–423 (1993)
4. Deb, K., Sundar, J., Rao, N.U.B., Chaudhuri, S.: Reference point based multiobjective optimization using evolutionary algorithms. *Int. J. Comput. Intell. Res.* **2**(3), 273–286 (2006)
5. Sanchis, J., Martínez, M.A., Ferragud, X.B.: Integrated multiobjective optimization and a priori preferences using genetic algorithms. *Inf. Sci.* **178**(4), 931–951 (2008)
6. Thiele, L., Miettinen, K., Korhonen, P.J., Luque, J.M.: A preference-based evolutionary algorithm for multi-objective optimization. *Evol. Comput.* **17**(3), 411–436 (2009)
7. Zitzler, E., Künzli, S.: Indicator-based selection in multiobjective search. In: Parallel Problem Solving from Nature, PPSN VIII, LNCS, vol. 3242, pp. 832–842 (2004)
8. Bader, J., Zitzler, E.: HypE: an algorithm for fast hypervolume-based many-objective optimization, Tech. Rep. TIK 286, Computer Engineering and Networks Laboratory, ETH Zurich (2008)
9. Bader, J., Zitzler, E.: Robustness in hypervolume-based multiobjective search, Tech. Rep. TIK 317, Computer Engineering and Networks Laboratory, ETH Zurich (2010)
10. Elhossini, A., Areibi, S., Dony, R.: Strength Pareto particle swarm optimization and hybrid EA-PSO for multi-objective optimization. *Evol. Comput.* **18**(1), 127–156 (2010)
11. Li, B.B., Wang, L.: A hybrid quantum-inspired genetic algorithm for multiobjective flow shop scheduling. *IEEE Trans. Syst. Man Cybern. Part B* **37**(3), 576–591 (2007)
12. Yang, D., Jiao, L., Gong, M.: Adaptive multi-objective optimization based on nondominated solutions. *Comput. Intell.* **25**(2), 84–108 (2009)
13. Ishibuchi, H., Murata, T.: A multiobjective genetic local search algorithm and its application to flowshop scheduling. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **28**(3), 392–403 (1998)
14. Lara, A., Sanchez, G., Coello, C.A., Schutze, O.: HCS: a new local search strategy for memetic multiobjective evolutionary algorithms. *IEEE Trans. Evol. Comput.* **14**(1), 112–132 (2010)
15. Goh, C.K., Tan, K.C., Liu, D.S., Chiam, S.C.: A competitive and cooperative co-evolutionary approach to multi-objective particle swarm optimization algorithm design. *Eur. J. Oper. Res.* **202**(1), 42–54 (2010)
16. Das, S., Mandal, D., Kar, R., Ghoshal, S.P.: A generalized closed form expression of directivity of arbitrary planar antenna arrays. *IEEE Trans. Antennas Propag.* **61**(7), 3909–3911 (2013)
17. Gozasht, F., Dadashzadeh, G.R., Nikmehr, S.: A comprehensive performance study of circular and hexagonal array geometries in the lms algorithm for smart antenna applications. *Prog. Electromagnet. Res. PIER* **68**, 281–296 (2007)

18. Goto, N.: Pattern synthesis of hexagonal planar arrays. *IEEE Trans. Antennas Propag.* **10**(8), 479–481 (1972)
19. Mahmoud, K.R., El-Adawy, M., Ibrahem, S.M.M., Bansal, R., Zainud-Deen, S.H.: A comparison between circular and hexagonal array geometries for smart antenna systems using particle swarm optimization algorithm. *Prog. Electromagnet. Res. PIER* **72**, 75–90 (2007)
20. Li, D.F., Gong, Z.L.: Design of hexagonal, planar arrays using genetic algorithms for performance improvement. In: Proceedings of 2nd International Conference on Microwave and Millimeter Wave Technology, pp. 455–460 (2000)
21. Deb, K.: Multi-objective optimization using evolutionary algorithms. Wiley, New York (2003)

QPSO for Synthesis of Scanned Linear Array Antenna for Fixed Side Lobe Level and First Null Beam Width Including Wide Null Placement

Hemant Patidar and Gautam Kumar Mahanti

Abstract In this research paper, quantum particle swarm optimization (QPSO) algorithm, an algorithm founded on the basic theory of particle cluster and properties of quantum mechanics, has been introduced for design of uniformly spaced scanned linear array antennas with definite value of sidelobe level (SLL), first null beam width (FNBW), and broad null depth. An effort is made such that these parameters are made equal to their related specific values. This is usually done by changing excitation current amplitude of the elements. The expression of array factor is derived by using the property of linear array antennas. One example has been presented with 35 isotropic antennas. Generated Pattern is scanned to an angle of 30°. The obtained results have shown that this technique is able to find required value of SLL, FNBW, and wide null depth. Although, the proposed method is developed and applied to a linear array of isotropic antennas; however, the principle can easily be applied to other types of arrays.

Keywords Linear array antenna · First null beam width · Quantum particle swarm optimization · Sidelobe level · Wide null depth

1 Introduction

Antenna array finds many applications in wireless communication system, radar, military, commercial, etc., where phased-array antennas are widely used in different applications, from military systems to commercial cellular communications networks. They are generally used in wireless communication systems to steer signals

H. Patidar (✉) · G.K. Mahanti
Department of Electronics and Communication Engineering,
National Institute of Technology, Durgapur 713209, India
e-mail: Hemantpatidar08@gmail.com

G.K. Mahanti
e-mail: gautammahanti@yahoo.com

from base stations to desired locations while creating nulls to suppress interference. Linear array is a very commonly used antenna array, which can constitute suitable narrow beam. Antenna arrays are intended to have a high directive beam; therefore, major part of the power should be transmitted by the main beam. This means that the array has to have a low radiation in the other directions. Nowadays, we are doing accrescent use of phased-array radar systems on naval ships and aircrafts. In a phased array, the elements are fed with non-uniform current excitations and phase shifters of all the elements are used to scan the beam in desired direction in free space. We can change the shape of pattern by controlling the array factor property by changing the geometrical configuration such as linear, planar, and rectangular and antenna parameters such as distance between elements, excitation of the elements, phase, and pattern of the individual elements [1, 2]. To fix the sidelobe level (SLL), first null beam width (FNBW), and wide null depth with scanned far-field pattern, elements are fed with non-uniform current excitation. But, the synthesis problem is more complex and difficult to solve with analytical methods. Several antenna array synthesis techniques were discussed in the letter [3–11].

To maximize the signal-to-noise ratio performance per undesirable interference, null steering system finds huge applications in radar and communication systems. It is achieved by placing the nulls along the direction of interference for antenna arrays [3–6]. This is done by synthesis of linear antenna array pattern with prescribed broad nulls [3]; genetic algorithm has been used for adaptive nulling by phase-only control method [4], a method for effects of null control on the radiation pattern is described in [5]. Null synthesis by controlling the phase for synthesis of antenna arrays [6] and nulling of a symmetrical pattern location by only the weight of phase control method was described in [7].

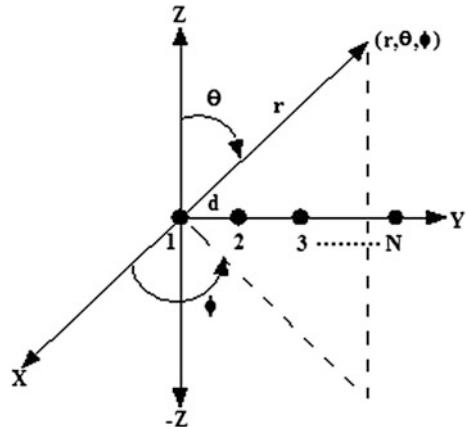
Here, the author discussed imposing of nulls in the desired directions for a scanned antenna radiation pattern. Comparison for phased-array synthesis by particle swarm optimization and genetic algorithms is described in [8]. A method for phase-differentiated reconfigurable array antenna including minimum dynamic range ratio was reported in [9]. The optimization relationship between beam width and SLL for broadside arrays [10] and a method for linear antenna arrays synthesis was discussed in [11].

In this research paper, the authors have used quantum particle swarm optimization (QPSO) algorithm [12–15] for synthesis of linear array of isotropic antennas.

2 Approach

An N -element linear array [1] is considered along the y -axis that is assumed uncoupled and uniformly spaced a distance d apart as shown in Fig. 1. Total generated field of the isotropic antenna array is calculated by vector summation of the field radiated by individual element. The free space [1] far-field pattern $\text{FP}(\theta)$ in the vertical plane is stated by (1):

Fig. 1 Geometry of an N -element linear array antenna along the y -axis



$$\text{FP}(\theta) = \sum_{n=1}^N B_n e^{i\phi_n} e^{i(n-1)kd \sin \theta} \quad (1)$$

where n = number of element, λ = the wavelength, B_n = the excitations amplitude, i = the imaginary unit, $k = 2\pi/\lambda$ = the wave number, d = spacing between the elements, $\phi_n = -(n - 1)kd \sin \theta_o$ = uniform progressive phase, θ_o = scan angle, and θ is the angle, represent the far-field point measured from z -axis as shown in Fig. 1.

Far-field in dB can be expressed in (2) as follows:

$$\text{FP}_n(\theta) = 20 \log_{10} \left[\frac{|\text{FP}(\theta)|}{|\text{FP}(\theta)|_{\max}} \right] \quad (2)$$

The main objective is to obtain the set of excitation current amplitude of the elements using QPSO that will satisfy our goal. The proposed algorithm (QPSO) will minimize the given cost function to obtain the desired value of SLL, FNBW, and wide null depth.

$$\text{Cost} = [vr_1 \times F_1^2 + vr_2 \times F_2^2 + vr_3 \times F_3^3] \quad (3)$$

where

$$F_1 = \begin{cases} \text{SLL}_o - \text{SLL}_d, & \text{if } \rightarrow \text{SLL}_o > \text{SLL}_d \\ 0, & \text{if } \rightarrow \text{SLL}_o \leq \text{SLL}_d \end{cases} \quad (4)$$

$$F_2 = \begin{cases} \text{FNBW}_o - \text{FNBW}_d, & \text{if } \rightarrow \text{FNBW}_o > \text{FNBW}_d \\ 0, & \text{if } \rightarrow \text{FNBW}_o \leq \text{FNBW}_d \end{cases} \quad (5)$$

$$F_3 = \begin{cases} \text{WN}_{\max}^o - \text{WN}_{\max}^d, & \text{if } \rightarrow \text{WN}_{\max}^o > \text{WN}_{\max}^d \\ 0, & \text{if } \rightarrow \text{WN}_{\max}^o \leq \text{WN}_{\max}^d \end{cases} \quad (6)$$

The coefficients vr_1 , vr_2 , and vr_3 determine the relative weight specified to each term in (3).

The value of $vr_1 = 15$, $vr_2 = 1$ and $vr_3 = 1$ has been assumed to generate the power pattern. SLL_d and SLL_o are desired and obtained values of sidelobe level, FNBW_o and FNBW_d are obtained and desired values of FNBW, and WN_{\max}^o and WN_{\max}^d are obtained and desired values of maximum wide null depth, respectively.

3 Quantum Particle Swarm Optimization (QPSO) Algorithm

QPSO algorithm [12–15] proposed by Sun et al. in 2004 is a novel optimization algorithm which is founded on the fundamental theory of particle swarm and properties of quantum mechanics in which all particles have the characteristic of quantum behavior. However, QPSO algorithm's global optimization performance is better than standard PSO algorithm because QPSO is stated only by the position vector; there is no velocity vector in QPSO. In QPSO, the dynamic behavior of the particle is different from that of the particle in standard PSO. According to uncertainty principle of quantum world, each particle can appear at any position of search space with certain probability, so the position and velocity of a particle cannot be found simultaneously.

The several steps involved in QPSO are given below:

- Step 1. Initialize a population of particles with random positions between the maximum and the minimum operating limits of the generators in the D -dimensional space.
- Step 2. Find out the cost value of each particle.
- Step 3. Now compare the personal best ($pbest$) of each particle with its present cost value. If the present cost value is better, then refer the present cost value to $pbest$ and refer the present coordinates to $pbest$ coordinates.
- Step 4. Calculate the mean best position ($mbest$) of all K particles using the following equation given below:

$$mbest = \frac{1}{K} \sum_{i=1}^K pbest_i$$

- Step 5. Determine the present best cost value in the entire population and its coordinates. Compare the global best position ($gbest$) of the current particle population with the previous one. If the present best cost value is

better than global best (g_{best}), then refer the current best fitness value to g_{best} and refer the current coordinates to g_{best} coordinates.

Step 6. Determine the vector local focus of the particle by the given equation:

$$f_{id}^t = \text{rand1}_{id}^t * (\text{pbest}_{id}) + (1 - \text{rand1}_{id}^t) * (\text{gbest})$$

Step 7. The d th dimension of the i th particle position (X_{id}) is updated by the following equations:

$$\begin{aligned} X_{id}^t &= f_{id}^t + (-1)^{\text{ceil}(0.5 + \text{rand2}_{id}^t)} \times \beta \\ &\quad \times |m_{best} - X_{id}^{t-1}| \times \log_e(1/\text{rand3}_{id}^t) \end{aligned}$$

If $X_{id}^t < X_{\min}^d$, then

$$X_{id}^t = X_{\min}^d + 0.25 * \text{rand4}_{id}^t * (X_{\max}^d - X_{\min}^d) \quad (7)$$

If $X_{id}^t > X_{\max}^d$, then

$$X_{id}^t = X_{\max}^d - 0.25 * \text{rand5}_{id}^t * (X_{\max}^d - X_{\min}^d) \quad (8)$$

where t is the current iteration, rand1 , rand2 , rand3 , rand4 , and rand5 are the uniformly distributed random numbers between 0 and 1. Equations (7) and (8) have been entered along all dimensions to squeeze the position within (X_{\max}^d, X_{\min}^d) if the particle tries to cross the desired domain of interest. These clipping techniques are sometimes essential to stop particles from eruption.

Step 8. The steps 2–7 above are repeated until a stop criterion is satisfied; ordinarily, the algorithm is stopped when there is no further updates in the best fitness value or reached maximum number of generation.

The parameter f is the local attractor of each particle. β is the contraction and expansion coefficient that control the convergence speed and performance of the particle. In order to avoid premature convergence, mean best position (m_{best}) is regarded as the barycenter of all particles. The value of β is 0.75. $\text{Ceil}(A)$ rounds the elements of A to the nearest integers greater than or equal to A .

4 Simulated Results

In this paper, a linear array of 35 isotropic antennas with element spacing of 0.5λ between the elements has been considered on the y -axis with specified SLL of -25 dB or less, specified FNBW of 10° or less and specified wide null depth (WN)

Table 1 Desired and obtained results

Design parameters	Desired value	Obtained value
Sidelobe level (dB)	-25	-25.12
First null beam width (°)	10	10.5
Maximum wide null depth in dB from ($\theta = 54^\circ$ to 61°)	-50	-50
Computation time (s)	...	1,393.54

of -50 dB or less from $\theta = 54^\circ$ to 61° to generate the free space [1] scanned far-field pattern $FP(\theta)$ at $\theta = 30^\circ$ in the vertical plane. Desired results are obtained by taking non-uniform current excitation amplitude of the elements for scanned linear antenna array.

QPSO is run for 600 iterations with a particle size of 60 to generate the far-field pattern in vertical plane. A set of 35 excitation current amplitude values is obtained.

The program is written in MATLAB for synthesis of the array antenna. The computation time is measured by a PC with Intel Core2 duo processor of clock frequency 2.93 GHz and 4 GB of RAM. Table 1 demonstrates the desired values and obtained results using QPSO. Figure 2 shows the best fitness value versus iteration number obtained by QPSO. Figure 3 shows the normalized current amplitude distribution to generate the free space far-field pattern with asymmetrical excitations. Figure 4 shows the normalized power pattern (far-field pattern) for scanned antenna array.

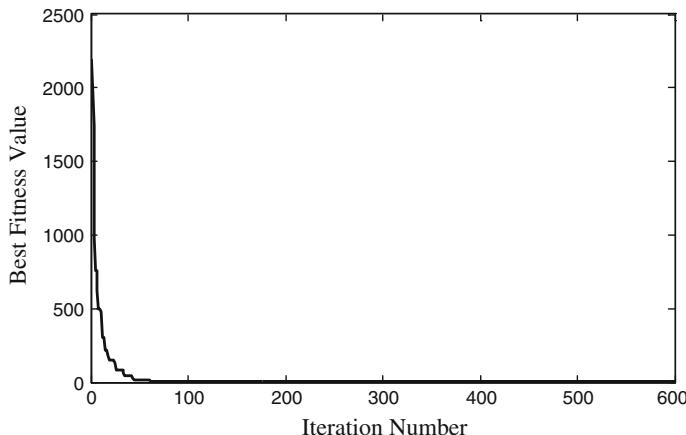
**Fig. 2** Best fitness value versus iteration number obtained by QPSO

Fig. 3 Distribution of normalized current excitation versus number of element obtained by QPSO

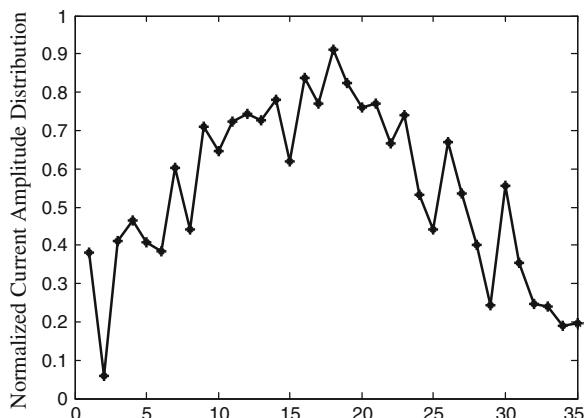
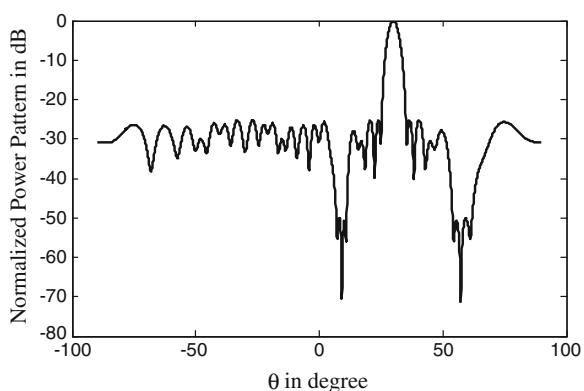


Fig. 4 Normalized far-field pattern in dB for linear array scanned at $\theta = 30^\circ$ using QPSO with prescribed wide null depth of -50 dB from $(\theta = 54^\circ \text{ to } 61^\circ)$



5 Conclusions

This paper used QPSO algorithm for synthesis of a scanned linear array antennas with uniform inter-element spacing to obtain fixed SLL, fixed FNBW, and fixed wide null depth. The obtained results using simulation depict a great deal between the desired and obtained specifications. The main benefit of using this technique is that it finds the suitable excitation current amplitude distribution using QPSO to obtain the desired values. Results obtained from synthesis of linear array antennas have illustrated the efficiency of this introduced technique.

References

1. Balanis, C.A.: Antenna Theory: Analysis and Design, 2nd ed. Wiley (Asia), Singapore (2003)
2. Haupt, R.L., Werner, D.H.: Genetic algorithms in electromagnetics. IEEE Press, New York (2007)

3. Er, M.H.: Linear antenna array pattern synthesis with prescribed broad nulls. *IEEE Trans. Antennas Propag.* **38**(9), 1496–1498 (1990)
4. Haupt, R.L.: Phase-only adaptive nulling with a genetic algorithm. *IEEE Trans. Antennas Propag.* **45**(6), 1009–1015 (1997)
5. Steyskal, H., Shore, R.A., Haupt, R.L.: Methods for null control and their effects on the radiation pattern. *IEEE Trans. Antennas Propag.* **34**(3), 404–409 (1986)
6. Vescovo, R.: Null synthesis by phase control for antenna arrays. *Electron. Lett.* **36**(3), 198–199 (2000)
7. Shore, R.A.: Nulling a symmetric pattern location with phase-only weight control. *IEEE Trans. Antennas Propag.* **32**(5), 530–533 (1984)
8. Boernerger, D.W., Werner, D.H.: Particle swarm optimization versus genetic algorithms for phased array synthesis. *IEEE Trans. Antennas Propag.* **52**(3), 771–779 (2004)
9. Mahanti, G.K., Chakraborty, A., Das, S.: Design of phase-differentiated reconfigurable array antennas with minimum dynamic range ratio. *IEEE Antennas Wirel. Propag. Lett.* **5**(1), 262–264 (2006)
10. Dolph, C.L.: A current distribution for broadside arrays which optimizes the relationship between beam width and sidelobe level. *Proc. IRE Waves Electrons* **34**, 335–348 (1946)
11. Murthy, P.K., Kumar, A.: Synthesis of linear antenna arrays. *IEEE Trans. Antennas Propagat.* **24**(6), 865–870 (1976)
12. Sun, J., Feng, B., Xu, W.: Particle swarm optimization with particles having quantum behavior. In: Proceedings of Congress on Evolutionary Computation, Portland, Oregon, USA, pp. 325–331(2004)
13. Sun, J., Xu, W.B., Feng, B.: A global search strategy of quantum behaved particle swarm optimization. In: Proceedings of IEEE Conference on Cybernetics and Intelligent Systems, Singapore, pp. 111–116 (2004)
14. Mikki, S.M., Kishk, A.A.: Quantum particle swarm optimization for electromagnetics. *IEEE Trans. Antennas Propag.* **54**(10), 2764–2775 (2006)
15. Ismail, T.H., Hamici, Z.M.: Array pattern synthesis using digital phase control by quantized particle swarm optimization. *IEEE Trans. Antennas Propag.* **58**(6), 2142–2145 (2010)

Decision Feedback Equalization for MIMO Systems

Samarendra Nath Sur, Rabindranath Bera and Bansibadan Maji

Abstract This paper deals with the implementation of decision feedback equalization-based receivers for MIMO systems in correlated Nakagami-m channel. Moreover, suboptimal, pre-equalization combining decision feedback receiver (DFE) architectures are examined and performance is compared with its linear counterparts.

Keywords MIMO · ZF · MMSE · DFE · Correlated channel

1 Introduction

The requirement for high performance in terms of secure communication and high-throughput communication system has led to the development of several signal processing and communications techniques for employing the resources efficiently [1, 2]. And to provide such a highly spectral efficient secure communication system, MIMO plays an important role [3]. The user demand of high data rate leads to intersymbol interference (ISI) and also non-orthogonality in the transmitted signal and that gives rise to multi-access interference (MAI) at the receiver. Therefore, in severe multipath scenario, the wireless system performance is severely degraded. Under these circumstances, a MIMO equalizer should be designed in a proper way to compensate for both intersymbol and multi-access interstream interference.

Now, to provide reliable and secure communications in such severe multipath channel condition, maximum-likelihood (ML) detection would be the optimum

S.N. Sur (✉) · R. Bera

Department of Electronics and Communication Engineering, Sikkim Manipal Institute of Technology, Sikkim, India
e-mail: samar.sur@gmail.com

B. Maji

Department of Electronics and Communication Engineering, National Institute of Technology, Durgapur, West Bengal, India

receiver. However, with the increase in the number of transmit antennas, the complexity of the receiver increases substantially [4]. This motivates the researchers to work on the minimization of the complexity of the receiver. Therefore, it is required to design low-complexity receiver such as linear receiver [5, 6] and decision feedback receiver (DFE) [7]. The DFE receiver has gained tremendous attention due to its improved performance and relatively low complexity as compared to ML receiver. Generally, in MIMO, zero forcing (ZF) and minimum mean square error (MMSE)-based receivers are used as linear receiver. In this paper, performance comparison between ZF and MMSE criterion-based linear (LE) and nonlinear (DFE) receiver has been carried out. Also in this paper, Cholesky factorization is utilized to design DFE receiver.

This paper presents the performance analysis of different MIMO receiver (linear and nonlinear) over the correlated Nakagami-m fading channel. Here, channel state information (CSI) is assumed to be present at the receiver side only. And it is also considered that the transmit power is uniformly distributed across the transmitting antennas. Moreover, Nakagami-m distribution provides a much realistic fading channel distribution and also it contains both the Rayleigh distribution ($m = 1$) and the uniform distribution on the unit circle ($m \rightarrow \infty$) as special (extreme) cases.

The paper is organized as follows: The system modeling is discussed in Sect. 2. We review the linear receiver ZF and MMSE algorithm and also DFE algorithm in the same section. Simulation results are presented in Sect. 3, and this paper is summarized in Sect. 4.

2 System Model

2.1 Channel Model

We consider a MIMO system [8], as in Fig. 1, with N_t transmitting antenna and N_r receiving antennas. The received signal y can be described by

$$y = Hx + n \quad (1)$$

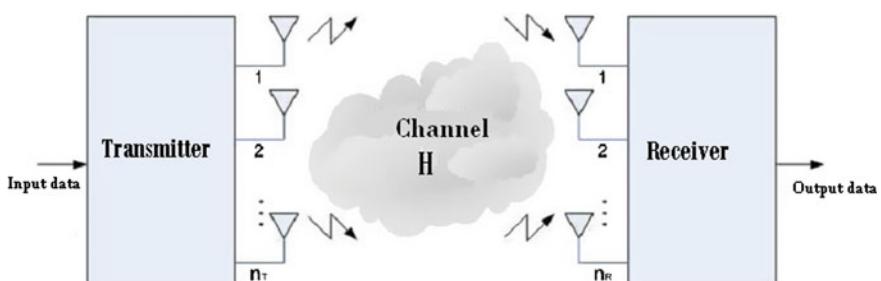


Fig. 1 MIMO system architecture

where the transmit symbol vector x satisfies $E\left\{\|x\|^2\right\} \leq P$ (P is the total power), and n is the $N_r \times 1$ additive white Gaussian noise vector. The vector H represents the correlated channels for the wireless transmission. The channel is assumed to follow a Nakagami-m fading probability distribution function (pdf). Let γ represent the instantaneous SNR, and it can be defined as

$$\gamma = \beta^2 \frac{E_s}{N_0} \quad (2)$$

where β is the fading amplitude, E_s is the energy per symbol, and N_0 is the noise spectral density. The probability distribution function of β for the Nakagami-m fading channel is given by

$$P_\beta(\beta) = \frac{2}{\Gamma(m)} \left(\frac{m}{\Omega} \right)^m \beta^{2m-1} \exp\left(-\frac{-m\beta^2}{\Omega}\right), \quad \beta \geq 0, \quad m \geq \frac{1}{2} \quad (3)$$

where $\Gamma(\cdot)$ represents the gamma function, $\Omega = E[\beta^2]$, and m is the parameter of fading depth that ranges from 0.5 to infinity, and this parameter is responsible for the variation in fading condition.

The fading parameter m is defined by the equation as given below

$$m = \frac{E^2[\beta^2]}{\text{var}[\beta^2]} \quad (4)$$

Then the pdf of the instantaneous SNR γ is given by [8]

$$P_\gamma(\gamma) = \frac{1}{\Gamma(m)} \left(\frac{m}{\bar{\gamma}} \right)^m \gamma^{m-1} \exp\left(-\frac{-m\gamma}{\bar{\gamma}}\right), \quad \gamma \geq 0 \quad (5)$$

2.2 Linear Receiver

The linear ZF and MMSE equalizers are important functional blocks and are ubiquitously required to design advanced receiver for MIMO system. The equalizer in MIMO system decouples the N_t transmitted data streams at the receiver.

It is a well-recognized fact that ZF receiver is a simple linear receiver, with low computational complexity. It minimizes interference but suffers from noise enhancement. ZF receiver works best with high SNR level. ZF implements matrix (pseudo)-inverse (+). The ZF receiver is

$$W_{\text{ZF}} = (H^H H)^{-1} H^H. \quad (6)$$

The MMSE equalizer (W_{mmse}) is obtained by minimizing the mean square error (MSE) defined as $E \left[\|x - W_{\text{mmse}}^H y\|^2 \right]$. The MMSE equalizer is given by

$$W_{\text{mmse}} = (H^H H + \rho^{-1} I)^{-1} H^H \quad (7)$$

MMSE detector helps to jointly minimize both the noise and interference or we can say that the MMSE detector seeks to balance between cancellation of the interference and reduction of noise enhancement. Therefore, MMSE detector outperforms the ZF detector in the presence of noise.

2.3 Pre-equalization Combining DFE

For the simplified architectural overview of DFE [9], the design of the zero-forcing (ZF) DFE is considered (Fig. 2).

As in the figure, first match filtering operation is performed on the received signal and can be mathematically expressed as given below,

$$p = H^* y \quad (8)$$

Then, p is passed through the feed-forward filter F and a feedback filter B to equalize. The coefficients for the above-mentioned two filters can be calculated using the Cholesky factorization of the equivalent channel matrix H_{equ} .

$$H_{\text{equ}} = H^* H = G^* \Gamma G \quad (9)$$

where Γ is diagonal with positive elements and G is upper triangular and monic. The feed-forward filter and the feedback filter are equal to

$$F = \Gamma^{-1} G^{-*} \quad (10)$$

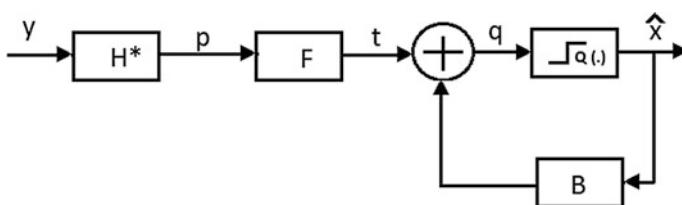


Fig. 2 DFE architecture

and

$$B = I - G \quad (11)$$

The estimate q of the desired symbols is given by

$$q = F^H p - B^H \hat{x} \quad (12)$$

where $\hat{x} = Q(F^H p)$ is the initial decision vector taken with the feed-forward section F and $Q(\cdot)$ represents a decision device.

The MMSE-DFE can be derived in an analogous fashion by performing Cholesky factorization on $H_{eq} + \sigma^2 I$.

3 Results

In this paper, the eigenvalue distributions of the correlation matrix of a MIMO channel with the spatial correlation of the multipath and with the variation of Nakagami-m fading parameter (m) are taken into account. The probability distribution of the largest eigenvalues of the correlation channel matrix is presented in this paper. The arrangement of the eigenvalues of the correlation matrix $H^H H$ is in descending order of $\lambda_1, \lambda_2, \dots, \lambda_m$ ($m = \min(N_t, N_r)$).

Figure 3 shows the variation in the eigenvalues with the variation in m values. As in the Fig. 3, the impact of m values is more on λ_1 in comparison with other values of λ . All the simulated results are extracted keeping SNR level at 0 dB.

Figure 4 shows the average values of the eigenvalues of $H^H H$ for different values of the correlation coefficient.

Fig. 3 Eigenvalues distribution for $N \times N$ MIMO

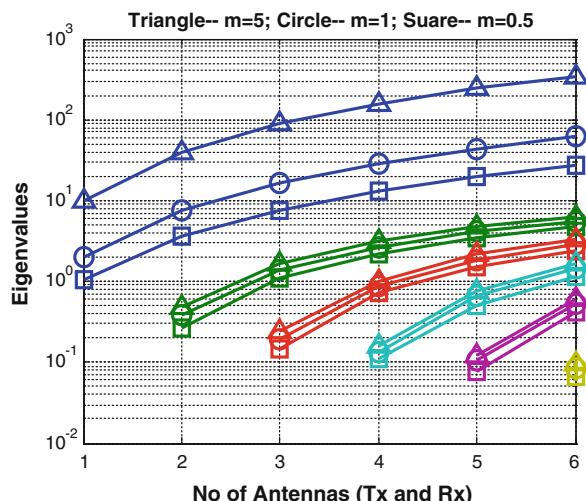
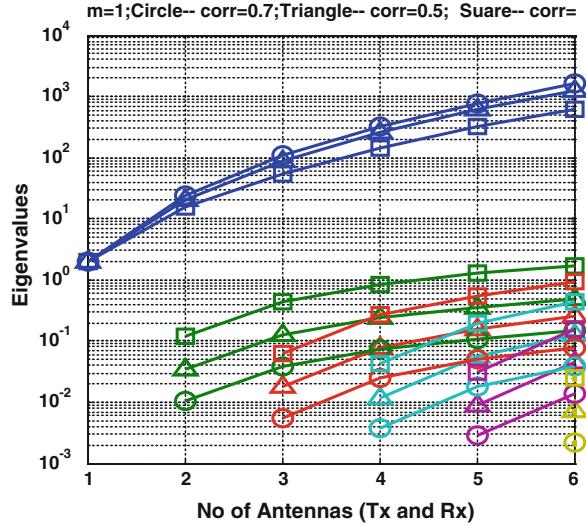


Fig. 4 Eigenvalues distribution for $N \times N$ MIMO



In this paper, the bit error rate (BER) is used for the comparison since it does not depend on the packet size of a specific system design. In a real system, the packet error rate should also be considered.

Figure 5 shows the performance comparison of a 2×2 MIMO system in Nakagami- m channel with QPSK modulation. As in figure, the performance improvement due to DFE receiver over LE receiver is clearly visible.

Figures 5, 6, and 7 represent the performance comparison of linear receiver like ZF, MMSE, and nonlinear receiver like ZF-DFE and MMSE-DFE. As depicted in figure, the improved (layered) optimal pre-equalization combining DFE that forms soft estimates based on all received symbols performs better than its linear

Fig. 5 Comparison between ZF and ZF-DFE

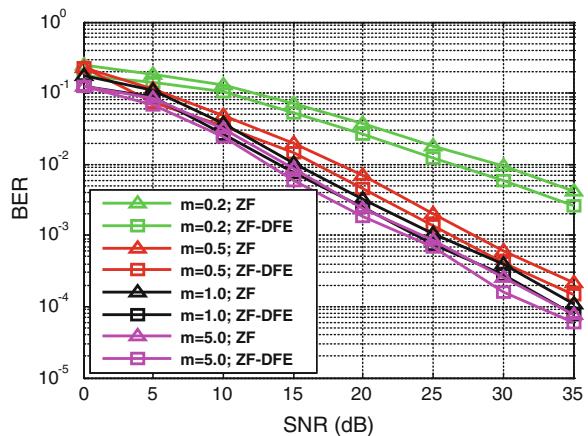


Fig. 6 Comparison between MMSE and MMSE-DFE

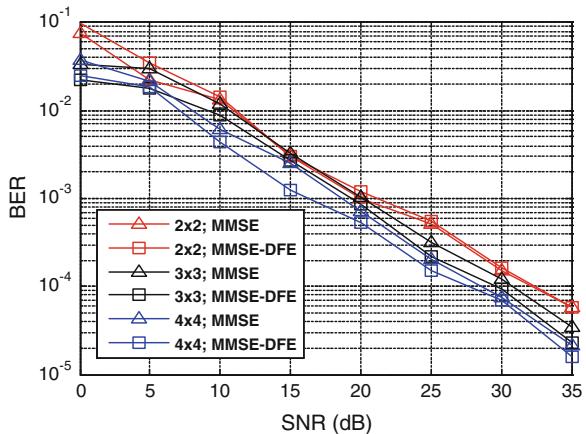
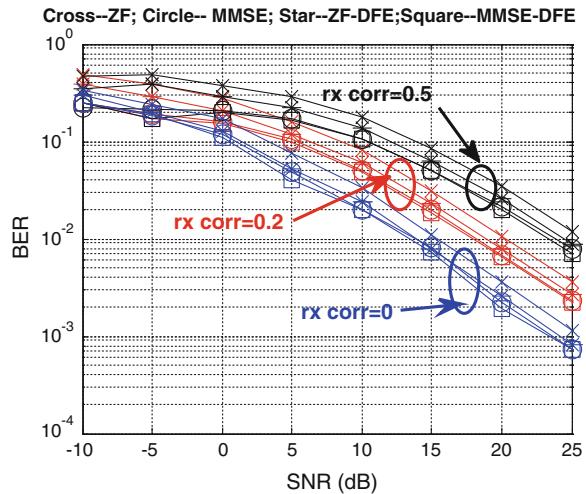


Fig. 7 Comparison between LF and DFE



counterpart. As represented in the simulated results, fading parameter m and channel correlation have significant influence over the MIMO system performance. In this severe channel condition, MMSE-DFE outperforms the other receiver.

4 Conclusion

This paper focused on the design of decision feedback equalization-based receivers for MIMO systems in correlated Nakagami-m channel. The equalizer coefficient matrices were derived using the Cholesky decomposition along MMSE/ZF

optimization. From the simulated results, it can be concluded that pre-equalization combining DFE achieves optimal performance and can be implemented to overcome severe channel condition.

References

1. Foschini, G.J., Gans, M.J.: On limits of wireless communications in a fading environment when using multiple antennas. *Wirel. Person. Commun.* **6**, 311–335 (1998)
2. Alamouti, S.: A simple transmit diversity technique for wireless communications. *IEEE J. Sel. Areas Commun.* **16**(8), 1451–1458 (1998)
3. Verd' u, S.: Minimum probability of error for asynchronous Gaussian multiple-access channels. *IEEE Trans. Inf. Theory* **32**(1), 85–96 (1986)
4. Duel-Hallen, A.: Equalizers for multiple input multiple output channels and PAM systems with cyclostationary input sequences. *IEEE J. Sel. Areas Commun.* **10**, 630–639 (1992)
5. de Lamare, R.C., Hjørungnes, A., Sampaio-Neto, R.: Adaptive decision feedback reduced-rank equalization based on joint iterative optimization of adaptive estimation algorithms for multi-antenna systems. In: IEEE WCNC, pp. 413–418 (2008)
6. Kominakis, C., Fragouli, C., Sayed, A.H., Wesel, R.: Multi-input multi-output fading channel tracking and equalization using Kalman estimation. *IEEE Trans. Signal Process.* **50**(5), 1065–1076 (2002)
7. Sur, S.N., Ghosh, D.: Channel capacity and BER performance analysis of MIMO system with linear receiver in Nakagami channel. *Int. J. Wirel. Microw. Technol.* **3**(1), 26–36 (2013)
8. Shin, H., Lee, J.H.: On the error probability of binary and M-ary signals in Nakagami-m fading channels. *IEEE Trans. Commun.* **52**(4), 536–539 (2004)
9. Lee, J., Toumpakaris, D., Jang, E.W., Lou, H.L.: DFE-based receiver implementation for MIMO systems employing hybrid ARQ. In: IEEE GLOBECOM, pp. 1–5 (2008)

Fuzzy Logic-Based Handover in 3GPP LTE Network

Parul Datta and Sakshi Kaushal

Abstract In today's world, people use Internet services and access applications on their mobile phones anywhere in the world. Wireless communication networks such as Third Generation Partnership Project's (3GPP's) Long Term Evolution (LTE) assist people to access Internet on their mobiles at high speed and in a seamless manner. When mobile devices move from a network from time to time, they need to be handed off to another network in order to provide users with same quality of service (QoS). In this paper, a FL-based handover scheme in LTE network is presented by considering parameters like received signal strength (RSS), data rate, and network coverage area. The handover scheme is implemented through simulation in NS-2. The results show improved packet delivery ratio (PDR) and decreased packet loss.

Keywords LTE · Fuzzy logic · Handover · RSS · PDR

1 Introduction

Wireless telecommunications networks are implemented to provide high-speed Internet access on mobile devices. For this purpose, many new technologies like Third Generation Partnership Project (3GPP) Long Term Evolution (LTE) were developed. Wireless network technologies are required to effectively exchange images and live video streams. LTE provides users with an all-IP solution at anytime and anywhere basis for data and voice traffic. Fourth-generation systems have two candidate systems: WiMAX and LTE [1]. LTE network architecture has two parts: E-UTRAN and EPC. Communications between the mobile nodes and the evolved

P. Datta (✉) · S. Kaushal
Computer Science and Engineering Department, UIET Panjab University,
Chandigarh, India
e-mail: paruldatta14@gmail.com

S. Kaushal
e-mail: sakshi@pu.ac.in

packet core (EPC) are handled by evolved universal terrestrial radio access network (E-UTRAN) [2]. E-UTRAN has only one component, i.e., evolved node B (eNB). The mobile nodes are controlled by eNB. The base station communicating with a mobile node will become its serving eNB. The EPC network has the following components: The Home Subscriber Server (HSS) is a central database containing information about subscribers of network operator. A default router for the user equipment (UE) and for communication with the outside world using SGi interface is packet data network gateway (P-GW). The data are forwarded between E-UTRAN and P-GW by serving gateway (S-GW). The high-level operations of managing mobiles and their sessions are controlled by mobility management entity (MME).

When an ongoing call or data session is transferred from one channel to another, it is known as handover. There is a possibility of handover in any technology [3]. The desirable features in handover process are as follows: It should be fast enough, low latency, QoS should be minimally affected, additional signaling during handover should be minimized, etc. Also, there are several handover performance metrics which should be considered for successful handover mechanism: handover call blocking probability, rate of handover, probability of unnecessary handover, etc. During handover, certain characteristics of a network like quality of service (QoS), throughput, performance of network, cost, and handover latency should be considered so as to satisfy user. Call dropping and latency should be minimized. Considering user needs, call termination, interference in signal, and several other handover failure reasons should be avoided as much as possible. In order to avoid handover failure, handover should be prioritized, i.e., different methods should be incorporated to handle and manage handover requests. Prioritization will avoid abrupt termination of ongoing calls.

The paper is organized as follows: Section 2 gives a brief insight into literature review. Section 3 gives FL-based handover algorithm. Section 4 presents the simulation analysis and results, and Sect. 5 concludes the paper.

2 Related Work

So far, a large number of researchers have discussed the handover issues by considering different QoS parameters. Some have considered received signal strength (RSS) and velocity of mobile terminal, while others considered RSS, data rate, network coverage area, etc., for effective handover management in LTE networks. Some of the schemes are discussed below.

LTE network has become the network technology choice for 4G deployments around the world. LTE's vision of wireless access has resulted in a comprehensive transition toward packet-switched only system as LTE represented significant shift from legacy mobile systems to all-IP network. LTE's all-IP architecture enables seamless delivery of applications and services. Authors in [4] have presented LTE specifications. Release 99 defined original dual-domain UMTS system supporting both circuit-switched voice services and packet-switched access. Release 4 was the

circuit-switched architecture which was bearer-independent. Release 5 through 7 increased the efficiency and improved data rates. Release 8 marked the start of transition to 4G technologies. Release 9 offered enhancements to LTE which included home evolved node base stations (eNBs) for improved residential coverage. Release 10 defined LTE-Advanced (LTE-A) which includes operational efficiencies by supporting self-optimizing, self-healing capabilities. Also, various LTE QoS Class Identities are discussed. Class identity represents QoS features a bearer offers. Network operators preconfigure class identity characteristics in each eNB. Each data path in LTE is assigned a set of QoS criteria. For users requiring different services, additional bearer paths are added. To harvest the benefits of LTE, network operators face many challenges such as migration strategies from legacy 2G/3G networks and deployment of IP networks to deliver low latency in order to support real-time QoS. Deploying LTE to meet such challenges will lead to improved network performance and overall cost savings across the network. LTE introduced IP-based architecture, with radio resource management and mobility management, where handover decisions are made by eNBs [5]. During intra-system handover, UEs are not connected to the system for some time. This is the time, i.e., detach time, when user traffic is being forwarded from source eNB to target eNB. Length of this time and some extra delay may lead a decrease in QoS. The traffic is prioritized based on the service: handover control messages having highest priority, VoIP messages having second highest priority, and data traffic having least priority. A strict priority scheduler was implemented in routers which follows the prioritization scheme. The priority scheme and alternative hierarchical scheduling architecture improved the performance of the system in terms of data rate, latency, and cost.

Efficient vertical handover algorithm needs to be designed which maintains network connection and provides acceptable resources to ongoing services is a challenging issue. Handover mechanism has three steps: Initiation is done by measuring RSS and mobile terminal velocity, the decision is taken using FL, and execution is done using Mobile IPv6. In [6], Next Generation Vertical Handover Decision Algorithm (NG-VDA) has been proposed to perform vertical handover between LTE and next generation wireless LAN. Parameters such as bandwidth, signal-to-noise ratio, life of the battery, and load on the network are considered. FL has been used to know most critical parameter for efficient system performance. FL system has three parts: Fuzzifier which converts crisp input into a fuzzy variable, the variables are used in rule base of inference engine and defuzzifier for converting fuzzy variable into a number called Handover Authorization Unit which takes final decision for handover. The proposed algorithm helped in achieving efficient vertical handover.

Reduction in handoff blocking probability is very important in order to improve QoS. Users get more frustrated when an ongoing call is dropped and not when a new call (NC) is denied admission [7]. A connection admission control (CAC) scheme is designed to reduce handoff dropping. Call admission scheme is used to prioritize handoff call (HC) over NC. A resource block is reserved and priority is given to handoff requests in admission. Three types of services are considered: VoIP, video streaming, and data on demand. On arrival of HC and NC calls to an overloaded cell, the NC will await acceptance by storing NCs in specific queues.

For acceptance of HC, a degradation procedure is triggered to degrade number of resource blocks (RBs) allocated to some ongoing calls and to allocate those RBs to HCs. When a call leaves the cell, total available bandwidth is increased. The proposed RB reservation algorithm uses increased or decreased RBs for calls in a cell. This algorithm achieved an improvement over call blocking probability, decreased dropping probability, and better resource utilization.

In [8], the authors have done in-depth comparison of LTE and WiMAX by considering different parameters. The comparison is divided into three parts: throughput comparison of downlink in time division duplexing (TDD); uplink TDD throughput results and LTE throughput in TDD; and frequency division duplexing (FDD) mode to understand LTE throughput in a better way. Performance of LTE is better in downlink and uplink. In [9], the authors have presented comparative study on network architecture and security of LTE and WiMAX networks. A research test bed was developed to gain insights into 4G wireless network in the enterprise, integration challenges, and key research problems. Both WiMAX and LTE have their own specified network architecture. WiMAX has BS, ASN gateway, AAA server, HA server, and some other components. LTE has eNB, S-GW, P-GW, MME, and HSS. In security architecture of WiMAX, authentication process uses either EAP_TTLS or EAP_TLS which uses private certificates and enterprise controlled username and password. In security architecture of LTE, authentication process uses EAP_AKA which requires new ways to integrate enterprise credentials. Authors concluded that both WiMAX and LTE can be deployed by enterprises as both provide high capacity, wide coverage range, QoS mechanisms, and security by using enterprise authentication.

For efficient handover decision making, more intelligent approaches need to be applied to improve the performance results for both the user and the network. The authors in [10] introduced a new adaptive vertical handover decision algorithm in which genetic algorithms are used to optimize fuzzy membership functions. The authors proposed a smart mobile terminal (SMT) to scan the environment for available radio access technologies (RATs). It evaluated the working condition of RATs using its newly developed hybrid fuzzy-genetic algorithm-based vertical handoff algorithm, triggers handoff process, if necessary, and chooses the best access point (AP) to camp on. Genetic algorithms are based on adaptive search techniques based on genetic rules to shape fuzzy membership functions. The proposed algorithm improved the performance by determining whether a handoff is required or not and noticeably reduced the number of handoffs. Managing mobility with seamless handoff and QoS guarantees is vital in 4G wireless networks. Mobility management enables location of roaming terminals to deliver data packets and maintaining connections when moving to new subnet. Current mobility management schemes have many disadvantages such as traffic overhead, high packet loss, high handoff latency, etc. In [11], the authors have proposed efficient handoff protocol for 4G called Handoff Protocol for Integrated Networks. The protocol alleviated service disruption. It also allowed seamless roaming and service continuity across various networks. The results showed improvement in performance as compared to existing schemes in terms of handoff blocking probability.

3 Fuzzy Logic-Based Handover Algorithm

A simple way to reach specific conclusion based upon unclear information is defined as fuzzy logic (FL) [12]. FL does not have any crisp limits. Range of FL variables lies between 0 and 1. Rule-based FL incorporates a simple rule-based approach to solve a problem. Fuzzy analysis follows three basic steps: fuzzification of inputs, fuzzy inference, and defuzzification for obtaining crisp output. Syntax of rule-based FL is shown in Eqs. (1) and (2).

$$\text{IF } x \text{ THEN } y \quad (1)$$

$$\text{IF } x \text{ and } y \text{ THEN } z \quad (2)$$

In this study, rule-based FL is considered to take decision whether to do handover or not based on RSS values measured by the UE. Three input parameters are considered, and fuzzy sets are assigned to each of them. Fuzzy set values for RSS, data rate, and network coverage area consists of high, medium, and low. There are three input parameters and three fuzzy sets for each input parameter, respectively. Hence, the maximum possible number of rules in our rule base is the following: $3^3 = 27$.

3.1 Fuzzy Rule Base

Table 1 presents the rules used in this study in which RSS, data rate, and network coverage area are checked for their values, i.e., high or medium or low. Based on the above rules, decision is taken whether to do handover or not. The RSS value indicates that how well the signal is being transmitted by the AP and is reaching the node [13]. On the other hand, RSS does not indicate how well the AP is hearing. RSS depends on a number of factors such as output power of transmitter, sensitivity of receiver, the gain of antennae at both ends of path, path loss, and attenuation of signal as it passes from transmitter to receiver. Sometimes due to low power levels and attenuation of free space, RSS value is expressed as a negative number. The more negative the number, the weaker the signal strength; on the other hand, if the number is closer to zero, then the signal strength is stronger.

The following steps helped in the handover process:

- Initially, UE is connected to a base station known as source base station. When UE experiences low-quality signal, it will measure RSS of source base station, i.e., the base station it is currently connected to and that of the target base station, i.e., the base station it will connect to in future.
- If the RSS of target base station is greater than that of the source base station, UE will handover itself to the target base station.
- RSS is calculated by the formula given below in Eq. (3) [14]:

Table 1 Fuzzy rule base

No.	(IF) RSS (is)	(and IF) data rate (is)	(and IF) network coverage area (is)	(THEN) HO decision (is)
1.	Low	Low	Low	Yes perform HO
2.	Low	Low	High	Probably yes
3.	Low	Medium	High	Probably yes
4.	Low	High	High	Probably yes
5.	Medium	Low	High	Probably no
6.	Medium	Medium	High	Probably no
7.	Medium	High	High	No
8.	High	Low	High	No
9.	High	Medium	High	No
10.	High	High	High	No
11.	Low	Low	Medium	Yes perform HO
12.	Low	Medium	Low	Probably yes
13.	Low	High	Medium	Probably yes
14.	Medium	Low	Low	Yes perform HO
15.	Medium	Medium	Medium	Probably no
16.	Medium	High	Low	Probably yes
17.	High	High	Low	Probably no
18.	High	Low	Medium	No
19.	High	Medium	Low	No
20.	High	High	Medium	No
21.	Low	Medium	Medium	Yes perform HO
22.	Low	High	Low	Yes perform HO
23.	Medium	Low	Medium	Probably yes
24.	Medium	Medium	Low	Probably no
25.	Medium	High	Medium	Yes perform HO
26.	High	Low	Low	Yes perform HO
27.	High	Medium	Medium	Probably no

$$\text{RSS} = \text{Pt} * \text{Gr} * \text{Gt} * \left(\frac{\pi * \pi * d * d}{16} \right) \quad (3)$$

where Pt is transmit power, Gt is antenna gain of transmitter, Gr is antenna gain of receiver, and d is the distance $\sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2}$.

Fuzzy handover decision is determined by using the formula [15] given below:

$$\text{Handover Probability} = \frac{\sum Mi * Wi}{\sum Mi} \quad (4)$$

where Mi is the membership degree defined as 0.3 for low, 0.5 for medium, and 1 for high and Wi is the handover probability weight value as considered from Table 2.

Table 2 Table for handover probability weight values

Value	Yes	Probably yes	No	Probably no
Weight	1	0.8	0.2	0.4

To determine when handover is required, the handover probability is computed and is used as follows: If handover probability >0.65 , then initiate handover otherwise do nothing. In this work, after applying the defuzzification step, handover probability is obtained as:

$$\text{Handover Probability} = \frac{0.3 * 0.8 + 0.5 * 0.8 + 0.3 * 0.8}{0.3 + 0.5 + 0.3} = 0.8$$

As the calculated value of handover probability, i.e. 0.8, is greater than the defined threshold of 0.65, handover will be initiated and UE will connect to target base station with better QoS parameters.

3.2 Algorithm

The algorithm describes handover process in which UE is initially connected to source eNB and is periodically measuring the RSS of source eNB. If UE experiences low RSS, then it will check data rate. If data rate is also low, then network coverage area is checked. If network coverage area is also low, then RSS of target eNB is measured. If RSS of target eNB is greater than that of source eNB, then perform handover. The data rate input parameter is defined from 0 to 55 Mbps. The network coverage area is defined from 0 to 30 km.

Handover Algorithm.

```

BEGIN
Step 1: UE connected to source eNB
Step 2: UE periodically measures RSS of source eNB
Step 3: If RSS is low go to step 3.1
    Else go to step 4
        Step 3.1: If data rate is low go to step 3.2
        Else go to step 4
            Step 3.2: If network coverage area is low then measure RSS of target
                base stations go to step 3.3
                Else go to step 4
                    Step 3.3: If  $\text{RSS}_{\text{source eNB}} < \text{RSS}_{\text{target eNB}}$  then perform hand-
                        over
                        Else go to step 4
Step 4: No need of handover
Step 5: UE connects to target eNB
END

```

4 Simulation Analysis and Results

4.1 Packet Delivery Ratio

Packet delivery ratio (PDR) is the ratio of number of delivered data packets to the destination to the number of packets that have been sent [16]. PDR is calculated by the formula given in Eq. (5). Figure 1 shows graph for PDR versus simulation time.

$$\text{PDR} = \frac{\sum(\text{Number of packets delivered})}{\sum(\text{Number of packets sent})} \quad (5)$$

The simulation trend seen in Fig. 1 for PDR depicts that more packets are reaching the destination after handover; hence, PDR improved by 30 %.

4.2 Packet Loss

Packet loss is the total number of packets dropped during simulation [16]. Packet loss is calculated by the formula given in Eq. (6). Decreased packet loss leads to better performance. Figure 2 shows graph for packet loss versus simulation time.

Fig. 1 Graph for packet delivery ratio versus simulation time

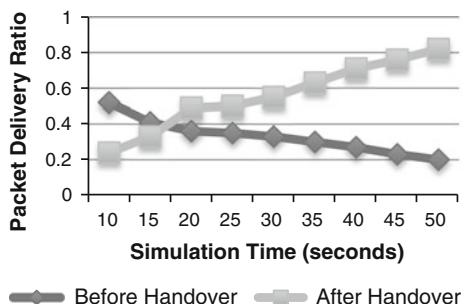
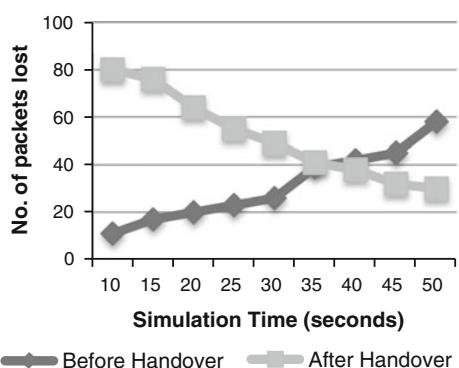


Fig. 2 Graph for packet loss versus simulation time



$$\text{Packet Loss} = \text{Number of packets sent} - \text{Number of packets lost} \quad (6)$$

The simulation trend seen in Fig. 2 for packet loss depicts that number of packets dropped after handover are less in comparison if no handover is performed; hence, packet loss decreased by 33 %.

5 Conclusion

In this paper, certain QoS parameters like PDR and packet loss are improved by performing handover. The simulation results show that PDR increased by 30 % after handover. There is a significant increase in PDR after handover. So, handover in this case is vital as it leads to an increase in PDR and hence, better network performance. The simulation results also show that packet loss decreased by 33 % after handover. Packet loss decreased significantly after UE is connected to target base station. Handover led to decrease in packet loss as reduction in packet loss is considered good for a network to perform better. Future works will focus on using artificial neural networks and genetic algorithms.

References

1. WiMAX vs LTE—Which is a Better 4G Technology. <http://thebestwirelessinternet.com/wimax-vs-lte.html>
2. Datta, P., Kaushal, S.: Exploration and comparison of different 4G technologies implementations: a survey. In: IEEE International Conference on Recent Advances in Engineering and Computational Sciences (RAECS), pp. 1–6, Chandigarh (2014)
3. Handover. <http://en.wikipedia.org/wiki/Handover>
4. LTE Networks: Evolution and Technology Overview. <http://www.tektronixcommunications.com/LTE>
5. Bajzik, L., Horvath, P., Korossy, L., Vulkan, C.: Impact of intra-LTE handover with forwarding on the user connections. In: 16th IST Mobile and Wireless Communications Summit, pp. 1–5, Budapest (2007)
6. Aziz, A., Rizvi S., Saad, N.M.: Fuzzy logic based vertical handover algorithm between LTE and WLAN. In: International Conference on Intelligent and Advanced Systems (ICIAS), pp. 1–4, Malaysia (2010)
7. Khitem, B.A., Zarai F., Kamoun, L.: Reducing handoff dropping probability in 3GPP LTE Network. In: Second International Conference on Communications and Networking (ComNet), pp. 1–8, Tozeur (2010)
8. Rezaei, F., Hempel M., Sharif, H.: A comprehensive performance analysis of LTE and mobile WiMAX. In: 8th International Wireless Communications and Mobile Computing Conference (IWCMC), pp. 939–944, Limassol (2012)
9. Yi, L., Miao, K., Liu, A.: A comparative study of WiMAX and LTE as the next generation mobile enterprise network. In: 13th International Conference on Advanced Communication Technology (ICACT), pp. 654–658, Seoul (2011)

10. Çalhan, A., Çeken, C.: An optimum vertical handoff decision algorithm based on adaptive fuzzy logic and genetic algorithm. *Wirel. Pers. Commun.* **64**(4), 647–664 (2012)
11. Makaya, C., Pierre, S.: Efficient handoff scheme for heterogeneous IPv6-based wireless networks. In: IEEE Wireless Communications and Networking Conference, pp. 3256–3326, Kowloon (2007)
12. Fuzzy Logic-An Introduction. http://www.seattlerobotics.org/encoder/mar98/fuz/fl_part1.html#INTRODUCTION
13. Veris Aerospond Wireless Sensors: Received Signal Strength Indicator (RSSI). http://www.veris.com/docs/whitePaper/vwp18_RSSI_RevA.pdf
14. Exp-1: Understand the pathloss prediction formula. <http://203.110.240.139/VLS5/Exp1/Theory/expt1-theory.pdf>
15. Bchini, T., Tabbane, N., Tabbane, S., Chaput, E., Beylot, A.L.: Fuzzy logic based layers 2 and 3 handovers in IEEE 802.16e network. *Comput. Comm.* **33**(18), 2224–2245 (2010)
16. Packet Delivery Ratio, Packet Lost, End to End Delay. <http://harrismare.net/2011/07/14/packet-delivery-ratio-packet-lost-end-to-end-delay/>

FPGA Implementation of Novel Discrete Phase-Locked Loop

N. Bharani dharan, M. Chinnathambi and S. Rajaram

Abstract This paper presents a novel discrete phase-locked loop (DPLL) based on FPGA. In this DPLL, phase offset of reference sinusoidal signal is followed at the output of NCO with no dead zone (i.e., full phase lock-in-range). All the signals and components used in this DPLL system are realized as discrete digital time components. Hilbert transform is used to generate analytic signal, and then CORDIC algorithm in vector mode is used for instantaneous phase detection. A novel simple algorithm is used to compute the phase offset difference between reference sinusoidal signal and NCO output's discrete sinusoidal values. Either LUT-based sine wave generator or CORDIC-based sine wave generator is used in NCO. The total system is implemented using Xilinx ISE and compared with existing DPLL block. Also, performance of phase lock time and phase lock-in-range is analyzed.

Keywords DPLL · NCO · Phase detector · CORDIC algorithm · FPGA implementation · Hilbert transform

1 Introduction

More number of applications are shifting from analog domain to digital domain due to the advantages of better accuracy, the possibility to use more complexity, and higher predictability. In 1930s, phase-locked loop (PLL) concepts begin to evolve. It has since been used in number of applications especially in communication

N. Bharani dharan (✉) · M. Chinnathambi · S. Rajaram

Department of Electronics and Communication, Thiagarajar College of Engineering,
Madurai, India

e-mail: bharani2410@gmail.com

M. Chinnathambi

e-mail: chinnaa92@gmail.com

S. Rajaram

e-mail: rajaram_siva@tce.edu

systems [1]. In digital domain, discrete phase-locked loop (DPLL) is widely used for digital circuit's clock synchronization and FPGA's clock synchronization.

From the need of a flexible offset local oscillator in heavy ion particle accelerator [2, 3], the requirement of our proposed system has been derived. The FPGA is an ideal way to build DPLLs since it combines the flexibility of microprocessors as well as the speed and computational power of application-specific integrated circuits (ASICs).

This paper is organized as follows: Sect. 2 presents the proposed design for the DPLL system, and Sect. 3 gives the implementation of the proposed DPLL system in detail. Then, in Sect. 4, the results from simulations are presented, followed by conclusion in Sect. 5.

2 Discrete PLL

A DPLL or all digital phase-locked loop (ADPLL) is one of the types of PLL in which all the components and signals being used are digital in nature. Thus, FPGA implementation is easily possible for this type of PLL. The existing structure of DPLL [4–6] uses a PI controller to behave as a loop filter with characteristics of a low-pass filter. But in our proposed DPLL, a simple equation is used instead of loop filter since loop filter is comparatively complex. It uses Hilbert transform and coordinate rotation digital computer (CORDIC) algorithm as phase detection system for both reference and NCO output signal.

The basic idea of this type of phase detection system is to use the concept of analytic signal [7] in order to obtain the phase information. Let $H\{x(n)\}$ denote the Hilbert transform of an input signal $x(n)$. Then, analytic signal is formed as follows:

$$A(n) = x(n) + jH\{x(n)\} \quad (1)$$

By using the CORDIC iterative equations [8], the desired mathematical function can be performed by properly setting the initial values. CORDIC in rotation mode is used for sine wave generation in NCO, with initial values of $X_0 = 1$, $Y_0 = 0$, and $Z_0 = \theta$. Then, at the end of iterations:

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = \begin{bmatrix} \cos \theta \\ \sin \theta \\ 0 \end{bmatrix} \quad (2)$$

CORDIC in vector mode is used for instantaneous phase detection by using analytic signal produced by Hilbert transform, with initial values of $X_0 = \text{Re}\{A(n)\}$, $Y_0 = \text{Im}\{A(n)\}$, and $Z_0 = 0$, which gives the following equation at the end of iterations:

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = \begin{bmatrix} P\sqrt{\operatorname{Re}\{A(n)\}^2 + \operatorname{Im}\{A(n)\}^2} \\ 0 \\ \arctan \frac{\operatorname{Im}\{A(n)\}}{\operatorname{Re}\{A(n)\}} \end{bmatrix} \quad (3)$$

The NCO can be constructed by using either LUT method or CORDIC algorithm in rotation mode. The accuracy of the sine wave generation is inversely proportional to the step size. Generally, sine wave values vary from -1 to 1 . But it is tough to realize those intermediate fractional values in FPGA, and so the discrete sine wave values are multiplied by 100 , which makes the sine values vary from -100 to 100 instead of -1 to 1 . Both LUT method and CORDIC method will generate discrete sine wave values with a step size of 15° and so totally 24 discrete values are needed to finish one full cycle, i.e., $\sin 0, \sin 15, \sin 30, \dots \sin 345$, giving the values of $0, 25, 50, \dots -25$ (when multiplied by 100). These values are stored in array with ‘LUT address,’ which vary from 0 to 23 . By incrementing this address one by one, discrete sine wave values are called from array.

3 System Implementation

All the word sizes are of 8 bits since all the values in this system will have values ranging from -100 to 100 . Among many methods to make Hilbert transform, approximation of the Hilbert transform by using a FIR filter method is chosen due to its simplicity. The computation of the coefficients was optimized to the desired frequency range with the least squares method using the MATLAB filter design toolbox. The decrease in order of Hilbert filter affects only the magnitude of filter output, and phase information is not affected. So the filter order is selected after iterating with various orders of filter and the minimum order which gives the better result is chosen. The implemented Hilbert FIR filter has an order of 10 and coefficient word size of 8 bits.

For six iteration stages, 0.611 is the constant K that should be multiplied with the output sine values. But to reduce the percentage of error, we made $X_i = 61$ ($0.611 \times 100 = 61$), $Y_i = 0$, and $Z_i = 0$. Here, θ is varied as $0, 15, 30, 45, \dots, 345, 0, 15 \dots$ And so the sine values are computed at each rising edge of clock for various θ values.

In Fig. 1, x is the first element of $X(n)$, y is the first element of $Y(n)$, k is the first element of $K(n)$, and m is the first element of $M(n)$. The first element from the sequence is easily taken by using clocked—if statements in VHDL. The Eqs. 1 and 2 present in Fig. 1 are computed by using $y1$ and $y2$ which are present in the same figure as shown below.

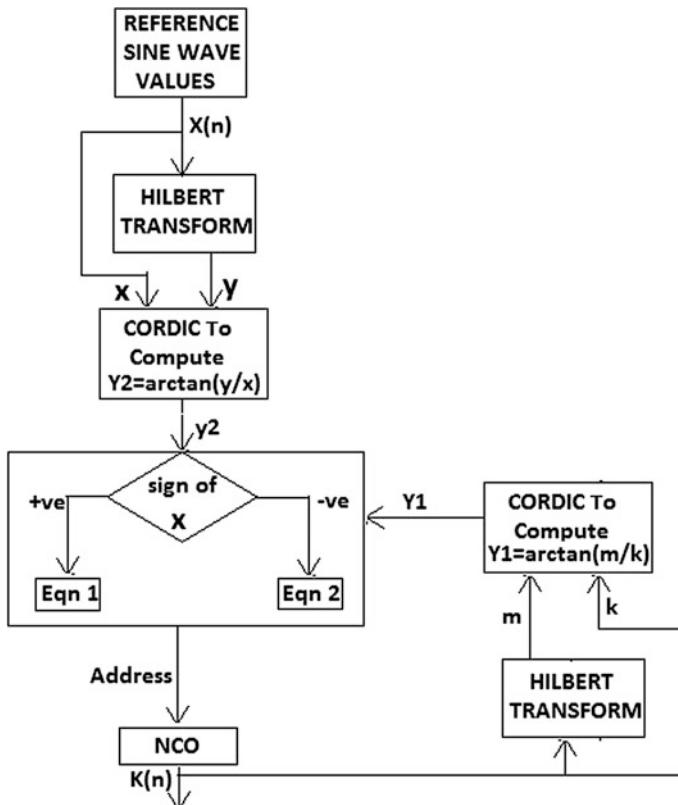


Fig. 1 Proposed structure of discrete PLL

For LUT-based sine wave generator

$$\text{Equation 1 : LUT address} = \frac{y_2 - y_1}{\text{stepsize}} + \frac{y_1 + 90}{\text{stepsize}}$$

$$\text{Equation 2 : LUT address} = \frac{y_2 - y_1}{\text{stepsize}} + \frac{y_1 + 90}{\text{stepsize}} + \frac{180}{\text{stepsize}}$$

For CORDIC-based sine wave generator

$$\text{Equation 1 : } Zi = (y_2 - y_1) + (90 + y_1)$$

$$\text{Equation 2 : } Zi = (y_2 - y_1) + (90 + y_1) + 180$$

The above simple novel equations compute the phase offset address required in NCO. Without y_1 and only with y_2 , it is possible to compute the phase offset

address. But y_1 is used to make general PLL structure with feedback. The values of y_1 and y_2 are also multiples of step size (here, step size = 15) varying as $-90, -75, \dots, -15, 0, 15, \dots, 75, 90, 75, \dots, 75$ for θ varying from $0, 15, 30, 45, \dots, 345$.

4 Results and Discussion

The hardware implementation of the DPLL has been done on a FPGA using Xilinx Spartan 3E board with xc3s500e-4fg320 as the target device in Xilinx ISE (Table 1).

In Table 2, the existing CORDIC block [4] which operates in vector mode is compared with total LUT- and CORDIC-based DPLL. It can be seen that in both methods, the area is very less, even when compared to existing single block of DPLL.

The limitation of using ordinary CORDIC algorithm is that the output will start to appear from $(N + 1)$ th clock pulse for N iteration stages. But after that, correct output will appear for each rising edge of clock cycle.

In Fig. 2, reference signal clock ('ref clk') is multiplied by 2 to form the NCO output clock ('output clock'). ' i ' represents the offset address value, and ' t ' represents the number of clock pulses. If the reference clock period is 10 ns, then 24 discrete values are required to make one complete sine wave cycle. Hence, time period will be 240 ns and so reference frequency will be 4.16 MHz. Also, NCO output clock period is 20 ns, time period will be 480 ns, and so NCO output frequency will be 2.08 MHz.

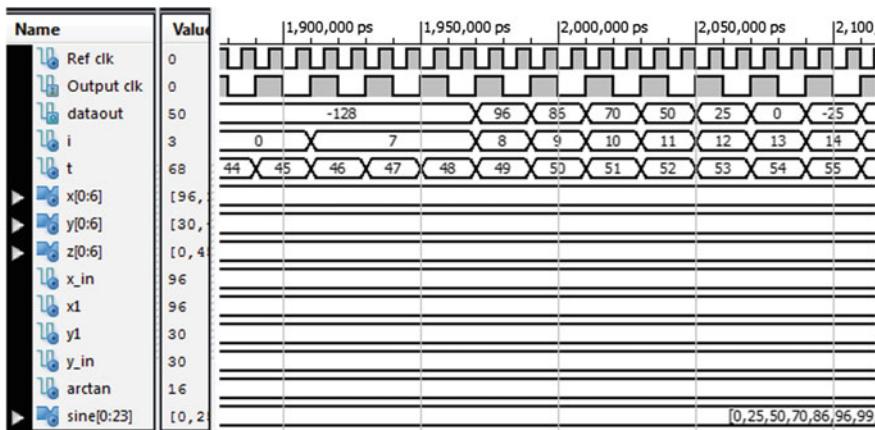
Thus, the NCO output frequency is divided by 2 and frequency synthesis operation is done (generally $F_{NCO} = M \times F_{ref}$). Here, the phase offset is 105° with value of $96[\sin(105^\circ)]$ at reference signal and it is followed at NCO output signal. Thus, phase offset correction is done at NCO. The DPLL output ('dataout') starts to appear from 49th clock pulse for LUT-based DPLL and 55th clock pulse for CORDIC-based DPLL. Hence, $0.96 \mu s$ ($20 \text{ ns} \times 48$) will be the phase lock time for the NCO frequency of 2.08 MHz in LUT method. Even if the NCO output frequency is 1 kHz, the phase lock time will be just 2 and 2.5 ms for LUT-based and CORDIC-based DPLL, respectively.

Table 1 Device utilization summary of various blocks of discrete PLL

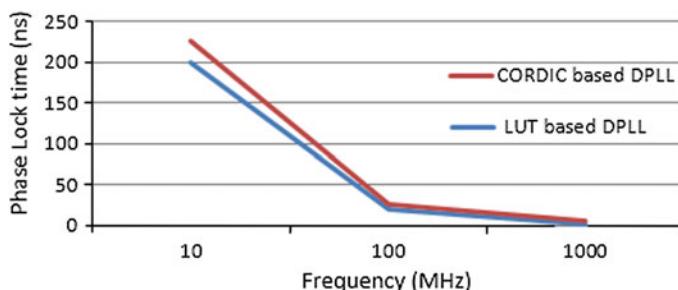
Logic utilization	Hilbert transform	CORDIC in vector mode	LUT-based NCO	CORDIC in rotation mode	Available
No. of slices	84	138	9	154	4,656
No. of slice flip-flops	155	136	14	143	9,312
No. of 4-input LUTs	99	262	14	302	9,312
No. of bonded IOBs	25	25	9	9	232
No. of GCLKs	1	1	1	1	24
No. of multipliers	7	—	—	—	20

Table 2 Device utilization summary comparison of proposed DPLL and existing CORDIC block

Logic utilization	LUT-based DPLL	CORDIC-based DPLL	Existing CORDIC block
No. of slices	275	420	592
No. of slice flip-flops	303	433	1,027
No. of 4-input LUTs	464	761	1,161
No. of bonded IOBs	26	26	66
No. of GCLKs	1	1	1
No. of multipliers	7	7	–

**Fig. 2** Simulation results of discrete PLL using LUT method

From Fig. 3, it is well known that phase lock time will decrease for increase in frequency. Performance analysis of DPLL with respect to phase lock-in-range is also made and it is verified for all the values of phase offset difference, and this DPLL is performing well without any dead zone for all the frequency range.

**Fig. 3** Performance analysis of DPLL with respect to phase lock time

5 Conclusion

In this paper, a detailed analysis of a simple and novel DPLL with full phase lock-in-range is made. From the device utilization summary, the proposed total DPLL system consumes very less area even when compared to single block of existing DPLL. Due to the usage of simple novel equation, DPLL with no dead zone became possible. But still, this novel equation does not reduce the accuracy of DPLL. The step size should be reduced from 15° , to give more accurate sine wave generation. But when step size is decreased more, CORDIC is suitable for sine wave generation as LUT needs more memory. Due to the changes we made in initial values as said in Sect. 3, we got percentage of error as only 2.42 % for CORDIC-based DPLL, while LUT-based DPLL is fully accurate. Also, the usage of DPLL in number of applications of digital communication systems is growing faster and this DPLL can still be modified and improvised upon in the near future to meet the needs of future demands for such designs.

References

1. Lata, K., Kumar, M.: ALL digital phase-locked loop (ADPLL): a survey. *Int. J. Future Comput. Commun.* **2**(6) (2013)
2. Klingbeil, H., Zipfel, B., Kumm, M., Moritz, P.: A digital beam-phase control system for heavy-ion synchrotrons. *IEEE Trans. Nucl. Sci.* **54**(6), 2604–2610 (2007)
3. Klingbeil, H.: A fast DSP-based phase-detector for closed-loop RF control in synchrotrons. *IEEE Trans. Instrum. Meas.* **54**(3), 1209–1213 (2005)
4. Das, A., Dash, S., Chitti Babu, B., Sahoo, A.K.: A novel phase detection system for linear all-digital phase locked loop. In: Proceedings of IEEE Students' Conference on Engineering and Systems (SCES 2012), pp. 1–6, MNNIT, Allahabad (2012)
5. Kumm, M., Klingbeil, H., Zipf, P.: An FPGA-based linear all-digital phase-locked loop. *IEEE Trans. Circuits Syst.-I Regul. Pap.* **57**(9), 2487–2497 (2010)
6. Das, A., Dash, S., Chitti Babu, B., Sahoo, A.K.: Design and implementation of FPGA based linear all digital phase-locked loop. In: India Conference (INDICON), 2012 Annual IEEE, pp. 280–285, Kochi (2012)
7. Kumm, M., Sanjari, M.S.: Digital Hilbert transformers for FPGA-based phase-locked loops. In: International Conference on Field Programmable Logic and Applications. FPL 2008, pp. 251–256 (2008)
8. Volder, J.E.: The CORDIC trigonometric computing technique. *IRE Trans. Electron. Comput.* **8**(3), 330–334 (1959)

Analytical Approach on the Scale Length Model for Tri-material Surrounding Gate Tunnel Field-Effect Transistors (TMSG-TFETs)

P. Vanitha, G. Lakshmi Priya, N.B. Balamurugan,
S. Theodore Chandra and S. Manikandan

Abstract In this paper, a new scale length theory for tri-material surrounding gate tunnel field-effect transistor (TMSG-TFET) has been proposed and derived. The scale length accounts for the effective conduction path of subthreshold leakage and thereby captures the short-channel effects (SCEs) and subthreshold factor. In order to derive the subthreshold swing in terms of scaling factor, the effective conducting path effect (ECPE) must be considered. Compared to conventional scaling theory, scaling of TMSG-TFET with ECPE has shown a much lower subthreshold slope (SS) of $S < 60$ mV/dec. The simulations of the proposed work are performed using 2D TCAD Sentaurus device simulator. The analytical results are compared and verified with the TCAD simulation results. Finally, results of the proposed work are compared with the scaling theory for MOSFETs with ECPE.

Keywords Scaling theory · Subthreshold behavior · Effective conducting path effect · Surface potential · Natural length · Scaling factor · Short-channel effects

P. Vanitha (✉) · G. Lakshmi Priya · N.B. Balamurugan · S. Theodore Chandra ·
S. Manikandan

Department of Electronics and Communication, Thiagarajar College of Engineering,
Madurai, India
e-mail: vani0104@gmail.com

G. Lakshmi Priya
e-mail: priya0217@gmail.com

N.B. Balamurugan
e-mail: nbbalamurugan@tce.edu

S. Theodore Chandra
e-mail: theodorechandra@gmail.com

S. Manikandan
e-mail: maniaugust36@gmail.com

1 Introduction

The traditional MOSFETs limit to certain problems such as high subthreshold slope (SS), high leakage current in off state, drain-induced barrier lowering (DIBL), and other short-channel effects (SCEs). Due to those limitations, scaling of MOSFETs below subthreshold swing 60 mV/dec is difficult. Among many devices, only TFET and I-MOSFET promise SS less than 60 mV/dec [1–4] and improved short-channel performance. So the promising device that will produce subthreshold swing below 60 mV/dec at room temperature is tunnel field-effect transistors (TFETs) due to the built in tunneling barrier.

The most effective way to increase the circuit performance is to reduce their size. As scaling technologies has drawn our attention toward reducing the channel length and increasing the performance of the device, it results in serious drawbacks of SCEs [5]. Scaling of the natural length gives a measure of the SCE inherent to a particular device structure [6]. A two-dimensional analysis is necessary to derive scaling theory that properly account for the channel length dependence. Yan et al. [7] proposed a unique scaling theory for DG MOSFETs by assuming that the punch-through current flows along the surface of the Si body. Suzuki et al. [8] proposed an alternative scaling to the length assuming that the maximum potential at the center of Si body is more sensitive to gate length than at the surface. Auth et al. [9] formulated a scaling theory for cylindrical, fully depleted, surrounding gate MOSFETs, and Chiang et al. [10] has proposed a new scaling theory for SG MOSFETs with effective conducting path effect (ECPE). Balamurugan et al. [11] has proposed a scaling theory for DMSG MOSFETs. Arun et al. [12] has proposed an analytical model for germanium single-gate silicon on insulator TFET.

Gate-engineered TMSG-TFETs are being explored in the recent research due to its advantage of lower SS of $S < 60$ mv/dec and improved carrier transport efficiency. TMSG-TFET device outperforms conventional TFET device due to the increased screening of the M1 region near the source from the drain potential variations and hence reduces the DIBL effect. Hence, tri-material surrounding gate tunnel FET device has been chosen for scale length model formulation.

In this paper, we propose a novel scaling theory for tri-material surrounding gate tunnel FET with ECPE. To derive scaling theory at subthreshold region, ECPE should be taken into account to improve the device performance and to suppress the SCEs. We derive a generalized scaling theory based on the new scaling factor α associated with ECPE. The analytical results are simulated using TCAD Sentaurus device simulator. With ECPE, our model helps us to reduce the SCEs and the effective minimum gate length for the proposed device is obtained.

2 Proposed Scaling Theory for Tri-Material Surrounding Gate Tunnel FET

The Fig. 1 represents the cross-sectional view of tri-material surrounding gate TFET with $L = 90$ nm as the length of the gate metal M . The gate metal M is made of three materials M_1 , M_2 , and M_3 of different work functions deposited over respective lengths L_1 , L_2 , and L_3 on the gate oxide layers. The device is composed of gate material 1 with the work function $\phi_{M_1} = 4.8$ eV (Au), gate material 2 with $\phi_{M_2} = 4.6$ eV (Mo), and gate material 3 with $\phi_{M_3} = 4.4$ eV (Ti).

2.1 Natural Length

The 2-D Poisson's equation for potential $\phi(r, z)$ in fully depleted TMSG-FETs is

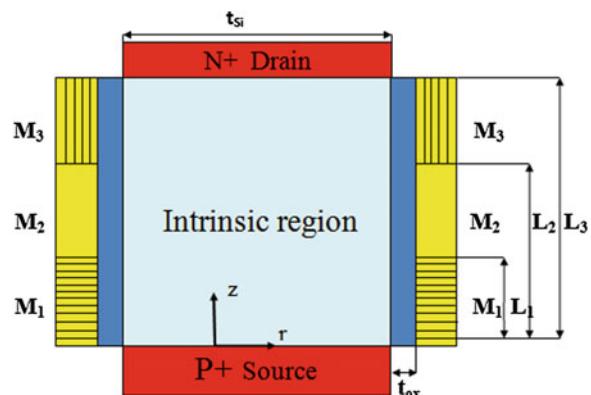
$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \phi(r, z)}{\partial r} \right) + \frac{\partial^2 \phi(r, z)}{\partial z^2} = \frac{qN_a}{\varepsilon_{si}} \quad (1)$$

where $\phi(r, z)$ is the electrostatic potential in the silicon film, q is the electronic charge, ε_{si} is the permittivity of the silicon film, N_a is the silicon film doping concentration. The potential profile in the vertical direction can be approximated by simple parabolic function as,

$$\phi(r, z) = s_1(z) + s_2(z)r + s_3(z)r^2 \quad (2)$$

The Poisson's equation is solved separately using boundary conditions.

Fig. 1 Cross-sectional view of tri-material surrounding gate tunnel FET



(a) The central potential is a function of z only.

$$\phi(0, z) = s_1(z) = \phi_c(z) \quad (3)$$

(b) The electric field at silicon/oxide interface is given by

$$\frac{\partial \phi(r, z)}{\partial r} \Big|_{r=R} = -\frac{\epsilon_{\text{ox}}}{\epsilon_{\text{si}}} \left(\frac{\phi_i(r, z) - \psi_{G_i}}{R \ln(1 + \frac{t_{\text{ox}}}{R})} \right) \quad (4)$$

(c) The electric field at $r = 0$ is equal to zero.

$$\frac{\partial \phi(r, z)}{\partial r} \Big|_{r=0} = 0 \quad (5)$$

where $\psi_{G_i} = V_{GS} - \phi_{m_i} + \chi + \frac{E_g}{2}$ where V_{GS} gate-to-source voltage, χ is the electron affinity of silicon, ϕ_{m_i} is the gate material work function, $i = 1, 2, 3$, E_g is the band gap energy, $2R = t_{\text{si}}$ is the diameter of the silicon pillar and t_{si} is the thickness of the silicon film, and ϵ_{ox} is the permittivity of the oxide layer.

On applying the boundary conditions (3) to (5) the potential in the effective conducting path can be obtained as,

$$\phi_{d_{\text{eff}}}(z) = \phi_c(z) + \left[\frac{\epsilon_{\text{ox}}}{\epsilon_{\text{si}}(2R)} \cdot \frac{\psi_{G_i} - \phi_s(z)}{\ln(1 + \frac{t_{\text{ox}}}{R})} \right] d_{\text{eff}}^2 \quad (6)$$

A simple scaling equation is obtained as,

$$\frac{d^2 \phi_{d_{\text{eff}}}(z)}{dz^2} + \frac{\psi_{G_i} - \phi_{d_{\text{eff}}}(z)}{\frac{1+B-Cd_{\text{eff}}^2}{4C}} = \frac{qN_a}{\epsilon_{\text{si}}} \frac{1-A}{1+B-Cd_{\text{eff}}^2} \quad (7)$$

Here, we define, $\lambda = \sqrt{\frac{1+B-Cd_{\text{eff}}^2}{4C}}$ in (7) as a natural length

$$\lambda = \sqrt{\frac{\epsilon_{\text{si}} 4R^2 \ln(1 + \frac{t_{\text{ox}}}{R}) t_{\text{si}} t_{\text{ox}}}{8\epsilon_{\text{ox}}} \left(1 + \frac{\epsilon_{\text{ox}}}{2\epsilon_{\text{si}} \ln(1 + \frac{t_{\text{ox}}}{R})} - \frac{\epsilon_{\text{ox}}}{\epsilon_{\text{si}}(2R^2) \ln(1 + \frac{t_{\text{ox}}}{R})} d_{\text{eff}}^2 \right)} \quad (8)$$

2.2 Scaling Factor

Once $\phi_{d_{\text{eff}}}(z)$ in (7) is determined, the potential of $\phi(r, z)$ can be uniquely solved by specifying two boundary conditions,

$$\phi_{d_{\text{eff}}}(z = 0) = V_{\text{bis}} \quad (9)$$

$$\phi_{d_{\text{eff}}}(z = L_{\text{eff}}) = V_{\text{bid}} + V_{\text{ds}} \quad (10)$$

where V_{bis} and V_{bid} are the built in potential of drain and source regions,

On solving Eq. (7) by using (9) and (10) and one obtains

$$\phi_{d_{\text{eff}}}(z) = \frac{1}{\text{Sinh}\left(\frac{L_{\text{eff}}}{\lambda}\right)} \left[(V_{\text{bi}} + V_{\text{ds}} + D) \text{Sinh}\left(\frac{z}{\lambda}\right) + (V_{\text{bi}} + D) \text{Sinh}\left(\frac{L_{\text{eff}} - z}{\lambda}\right) \right] - D \quad (11)$$

The subthreshold conduction can be analyzed by the minimum potential in the effective conducting path. The minimum potential with ECPE is given by,

$$\phi_{d_{\text{eff,min}}} \approx \sqrt{LM} \exp\left(\frac{-L_{\text{eff}}}{2\lambda}\right) - D \quad (12)$$

The scaling factor with ECPE is given as,

$$\alpha = \frac{L_{\text{eff}}}{2\lambda} \quad (13)$$

Thus, the novel scaling factor accounting for ECPE is derived. Once α is determined, the relationship between device parameter of t_{ox} and t_{si} can be expressed by

$$t_{\text{ox}} = \frac{\varepsilon_{\text{ox}}}{\varepsilon_{\text{si}}} \left[\frac{L_{\text{eff}}^2}{2\alpha^2 R} - \frac{R}{4} + \frac{d_{\text{eff}}^2}{2R} \right] \quad (14)$$

Equation (14) is the key equation to adjust the device parameters for allowable scaling factor α . Our model is more suitable for TMTFET's (Tri Material Tunnel Field Effect Transistor) scaling due to the third term on the right-hand side in (14) which takes various effective conducting paths corresponding to different doping densities into consideration, when compared to the model proposed by Auth et al.

3 Results and Discussions

Figure 2 shows the variation of minimum gate length with silicon thickness for various depths of effective conducting path. The minimum effective gate length is achieved for $d_{\text{eff}} = 0.55t_{\text{si}}$. With oxide thickness $t_{\text{ox}} = 3$ nm, a stable subthreshold swing of $S < 60$ mv/dec is achieved without a greater effective gate length.

Figure 3 depicts the dependence of subthreshold swing on scaling factor. It is clearly observed that as the natural length λ decreases, ideal subthreshold swing of $S < 60$ mV/dec is obtained, and hence, leakage current is also minimized. This ideal

Fig. 2 Variation of the minimum gate length with silicon thickness for various depths of effective conducting path

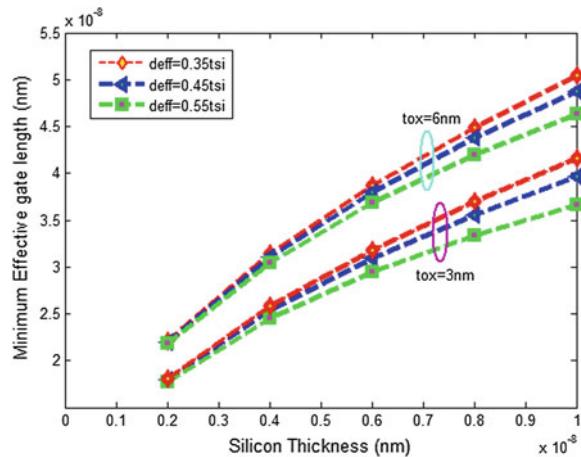
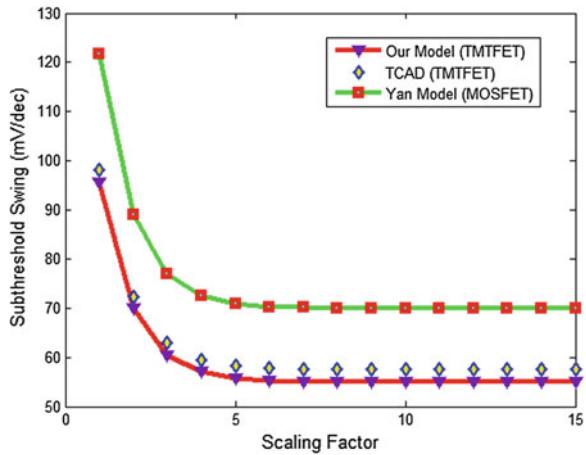


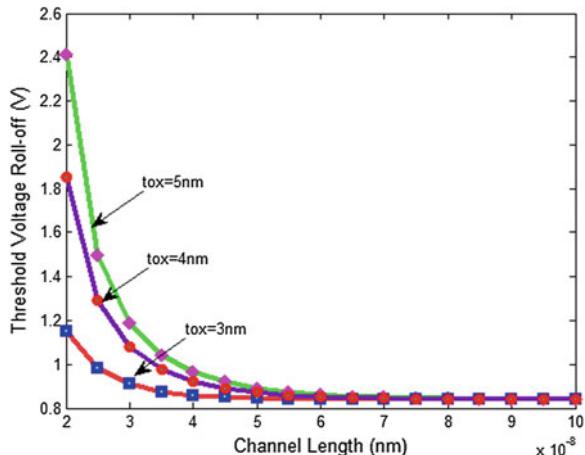
Fig. 3 Dependence of subthreshold swing on scaling factor and the analytical results are compared with simulation results



subthreshold swing, however, makes the scaling factor to go beyond the value of 5 and effectively suppresses the SCEs. When compared to MOSFETs, TMSG-TFETs provides a less SS. The analytical results of TMSG-TFETs are compared with TCAD simulation results, and a close match is observed with analytical and simulation results.

Figure 4 shows the channel length verses threshold voltage roll-off for various values of oxide thickness. The thin gate oxide of $t_{ox} = 3$ nm provides small threshold voltage roll-off, and it is clearly observed that this thin gate oxide greatly suppresses the SCEs. It is obvious that the thin gate oxide can be used in designing the TMTFET for small device technology due to the small threshold voltage roll-off.

Fig. 4 Variation of threshold voltage roll-off with channel length for various oxide thicknesses



4 Conclusion

ECPE is considered to play an important role in scaling theory. With ECPE, a novel scaling theory for tri-material TMSG-TFET is derived. The potential in the effective conducting path and a simple scaling equation has been derived, and on solving this equation, the minimum channel potential $\Phi_{d_{\text{eff}},\min}$ and the new scaling factor α is obtained to model the subthreshold behavior of high electron mobility transistors. The small natural length corresponds to good SCE immunity. The obtained analytical results are compared and verified against TCAD simulation results. The tri-material TMSG-TFET scaling theory with ECPE results is also compared with the scaling theory of MOSFET results, and it is well observed that due to ECPE, the device performance is increased in the subthreshold region, and it greatly eliminates the DIBL effects.

References

1. Yang, B., Buddharaju, K.D., Teo, S.H.G., Singh, N., Lo, G.Q., Kwong, D.L.: Vertical silicon-nanowire formation and gate-all-around MOSFET. *IEEE Electron Device Lett.* **29**(7), 791–794 (2008)
2. Yu, B., Chang, L., Ahmed, S., Wang, H., Bell, S., Yang, C.Y., Tabery, C., Ho, C., Xiang, Q., King, T.J., Bokor, J., Hu, C., Lin, M.R., Kyser, D.: FinFET scaling to 10 nm gate length. In: IEDM Technical Digest, pp. 251–254 (2002)
3. Gopalakrishnan, K., Griffin, P.B., Plummer, J.D.: I-MOS: a novel semiconductor device with a subthreshold slope lower than kT/q . In: IEDM Technology Digestivas, pp. 289–292 (2002)
4. Moselund, K.E., Ghoneim, H., Bjork, M.T., Schmid, H., Karg, S., Lortscher, E., Riess, W., Riel, H.: Comparison of VLS grown Si NW tunnel FETs with different gate stacks. In: Proceedings of the European Solid State Device Research Conference, pp. 448–451, Sep 2009

5. Kumar, S.P., Agrawal, A., Chaujar, R., Kabra, S., Gupta, M., Gupta, R.S.: Threshold voltage model for small geometry AlGaN/GaN HE MTs based on analytical solution of 3-D Poisson's equation. *Microelectron. J.* **38**(101), 3–1020 (2007)
6. Chaudhry, A., Kumar, M.J.: Controlling short-channel effect in deep-submicron SOI MOSFETs for improved reliability: a review. *IEEE Trans. Electron Devices* **4**(1), 99–109 (2006)
7. Yan, R.H., Ourmazd, A., Lee, K.F.: Scaling the Si MOSFET: from bulk to SOI to bulk. *IEEE Trans. Electron Devices* **39**(7), 1704–1710 (1992)
8. Suzuki, K., Tanaka, T., Tosaka, Y., Horie, H., Arimoto, Y.: Scaling theory for double gate SOI MOSFETs. *IEEE Trans. Electron Devices* **40**(12), 2326–2329 (1993)
9. Auth, C.P., Plummer, J.D.: Scaling theory for cylindrical, fully-depleted, surrounding-gate MOSFETs. *IEEE Trans. Electron Devices* **18**(2), 74–76 (1997)
10. Chiang, T.K.: A scaling theory for fully-depleted, surrounding-gate MOSFET's: including effective conducting path effect. *Microelectron. Eng.* **77**, 175–183 (2005)
11. Balamurugan, N.B., Sankaranarayanan, K., Suguna, M.: A new scaling theory for the effective conducting path effect of dual material surrounding gate nano scale MOSFETs. *J. Semicond. Technol. Sci.* **8**(1), 92–97 (2008)
12. Samuel, T.S.A., Balamurugan, N.B.: Analytical modeling and simulation of germanium single gate silicon on insulator TFET. *J. Semicond.* **35**(3), 034002 (2014)

Color Image Segmentation Using Cauchy-Mutated PSO

Suman Banerjee, Debasree Saha and Nanda Dulal Jana

Abstract Image segmentation is the process of subdividing a digital image into its constitute regions such that pixels belong to the same region will be same based on some image property (such as grayscale value, color, texture) and pixels in the different group will be different based on the same image property. Till date, different researchers have taken image segmentation problem from a different point of view and developed several image segmentation algorithms. This paper is going to address an optimization-based approach in color image segmentation where optimized threshold value is chosen by maximizing the Kapur's Entropy Function.

Keywords Particle swarm optimization · Threshold value · Color image · Cauchy-Mutated PSO

1 Introduction

Image segmentation [1, 2] is the process of grouping of pixels based on some homogeneity measure. It is one of the important steps in any computer vision and high-level image understanding applications. Actually, image segmentation is done to find out the region of interest and to reduce the data for further processing. Image segmentation can be applied in many real-life applications, content-based image retrieval [3], medical imaging [4], pattern extraction from an image [5], image compression [6], moving object detection [7], and different biometric applications [8]. A large amount of

S. Banerjee (✉) · D. Saha · N.D. Jana

Department of Information Technology, National Institute of Technology Durgapur,
Durgapur, West Bengal, India

e-mail: banerjeesuman1991@gmail.com

D. Saha

e-mail: debasreesaha01@gmail.com

N.D. Jana

e-mail: nanda.jana@gmail.com

literatures on different image segmentation algorithms exist. Among them, optimization-based approaches are quite recent. Different image segmentation techniques can be mainly categorized into two classes: discontinuity-based approaches and similarity-based approaches. Different image segmentation techniques include:

1. Edge/line/point detection.
2. Region-based segmentation.
3. Gray-level thresholding.
4. Clustering-based approaches.
5. Graph theoretic approaches.
6. Optimization-based approaches.
7. Entropy-based segmentation.

Remaining portion of this paper is organized as follows: Sect. 2 presents image segmentation in detail, Sect. 3 presents adaptive mutation-based PSO, Sect. 4 presents proposed algorithm, and Sect. 5 presents simulation and results. At last, Sect. 6 draws conclusions and gives future directions of this paper.

2 Detail about Image Segmentation

Image segmentation is the process of classifying the pixels of a digital image into its regions such that the following five conditions are satisfied:

1. $\bigcup_{i=1}^n R_i = I$ (Total Image Space).
2. R_i is a connected set.
3. $R_i \cap R_j = \emptyset$ for all i and j such that $i \neq j$.
4. $Q(R_i) = \text{True}$ for $i = 1, 2, \dots, n$.
5. $Q(R_i \cup R_j) = \text{False}$ for arbitrary two adjacent regions R_i and R_j .

where I is the entire image space and the entire image is divided into n regions R_1, R_2, \dots, R_n . The first condition says that the segmentation should be complete. The second condition says that segmented image is connected set. The third condition says that segmentation cannot be overlapping (though this condition can be relaxed for fuzzy segmentation). The fourth condition tells that the predicate(s) by which we are segmenting the image should be satisfied by all the pixels in that region, and the fifth condition tells that the same predicate cannot be satisfied simultaneously by more than one region.

3 Cauchy Mutation-Based PSO

Particle swarm optimization [9] is a swarm intelligence-based global optimization algorithm which mimics food-searching behavior of a birds' flock. Till date, different engineering design and optimization problems was solved by PSO algorithm.

Different researchers have proposed modified versions of this algorithm either by making parameters dynamic or by incorporating operators from some other algorithms. Each particle is a logical entity which flies over the search space, and each particle has two attributes Position (X) and Velocity (V). Particle's Position (X) represents a candidate solution in the search space. These attributes are updated generation after generation with the following formula:

$$V_i(t+1) = \omega * V_i(t) + c_1 r_1 (X_{\text{pbesti}}(t) - X_i(t)) + c_2 r_2 (X_{\text{gbesti}}(t) - X_i(t)) \quad (1)$$

$$X_i(t+1) = X_i(t) + V_i(t+1) \quad (2)$$

where the constant values ω , c_1 , and c_2 play an important role for the performance of this algorithm. Depending upon the neighborhood operator being used, PSO algorithm can be classified into two variants: *l-best* PSO and *g-best* PSO. Both the versions have got certain advantages as well as disadvantages. For *l-best* PSO, convergence rate is high but high probable to stuck at local optima and for *g-best* PSO it less probable to stuck at local optima but convergence rate is very slow. Hence, our goal is to use *l-best* PSO for its fast convergence rate, but we have to take corrective measures such that search process should not stuck at local optimum. So we have to use mutation operator which shifts the solution from local optima and also helps in exploration. Here, we have used *Cauchy Mutation Operator* for mutation purpose.

Algorithm 1: Fitness Function (Kapur's Entropy Function)

```

Input: Randomly initialized particles position( $X$ )
Output: Optimized Threshold Value
begin
    1. Read Input Image.
    2. Evaluate matrices in three different colorplanes.
    3. Calculate Probability( $P_i$ ) of each level(0 – 255) in different color planes.
    4. Calculate total probability for three different color planes by the following equations:
         $W_0 = \sum_{i=0}^{T-1} P_i$  and  $W_1 = \sum_{i=T}^{L-1} P_i$ 
    5. Calculate  $H_{ob} = \sum_{i=0}^{T-1} (P_i/W_0) \log_2 (P_i/W_0)$  and
         $H_{bg} = \sum_{i=T}^{L-1} (P_i/W_1) \log_2 (P_i/W_1)$ 
    6.  $H(T) = H_{ob} + H_{bg}$ 
    7. Return effective threshold  $H(T) = \sqrt{H(T_R)^2 + H(T_G)^2 + H(T_B)^2}$ 
end

```

4 Basics of Color Image and Proposed Algorithm

Color image [1] segmentation is an important step in color image processing. There exist different color models such as *R-G-B*, *C-M-Y*, and *H-S-V*. In RGB color model, a color image consists of three different color planes, one for each primary color *Red*, *Green*, and *Blue*. During the segmentation process, optimized threshold value is chosen in three different color planes. Our proposed algorithm

Cauchy-Mutated PSO is described in *Algorithm 2*, and fitness function (i.e., *Kapur's Entropy Function*) [10] is described in *Algorithm 1*. In [11], a noble approach for segmenting a grayscale image by *Kapur's Entropy Function* by optimizing PSO has been described. We have extended that work for color images and applied *Cauchy Mutation-Based PSO* [12] instead of simple PSO.

5 Simulation and Result

We have run our algorithm for four different color images with the following parameter settings: $\omega = 0.732$, $c_1 = c_2 = 1.4896$, and the optimized threshold value is shown in the following table:

Algorithm 2: Cauchy Mutated PSO

```

Input: Randomly initialized particle's position  $X(i)$  and velocity  $V(i)$  with dimension D
        and constant values  $\omega = 0.732$ ,  $C_1=C_2=1.4962$  and  $G=2000$ 
Output: Current Best Solution.
begin
    for  $k = 1$  to  $G$  do
        for  $i = 1$  to  $N\_POP$  do
            for  $j = 1$  to  $D\_POP$  do
                Update Particle's Position and Velocity by equation_no 1 and 2.
                 $Fitness(i) = Func(X(i, j), D\_POP)$ 
                if  $Fitness(i) \leq Local\_Best(i)$  then
                     $X_{best}(i) = Fitness(i)$ 
                     $Min\_best = min(Local\_best)$ 
                if  $Min\_best \leq Global\_Best$  then
                     $Global\_Best = min\_best$ 
            for  $i = 1$  to  $N\_POP$  do
                 $cauchy.vel.new(i) = V(i) * exp(trnd(D\_POP, 1, 1))$ 
                 $cauchy.pos.new(., i) = X(., i) + cauchy.vel.new(i, :) * trnd(D\_POP, 1, 1)$ 
                 $cauchy.Fitness(i) = func(cauchy.pos.new(i, :))$ 
                if  $Fitness(i, 1) \geq cauchyFitness(i)$  then
                     $X(i) = cauchy.pos.new(i)$ 
                     $V(i) = cauchy.pos.new(i)$ 
    end

```

Image name	$H(T) = \sqrt{H(T_R)^2 + H(T_G)^2 + H(T_B)^2}$
Byke.jpg	37
Baboon.jpg	175
Pepper.jpg	125
Airplane.jpg	174

Firstly, original images with their histograms of all the three planes are shown, and then their corresponding segmented images are shown in Figs. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, and 20.

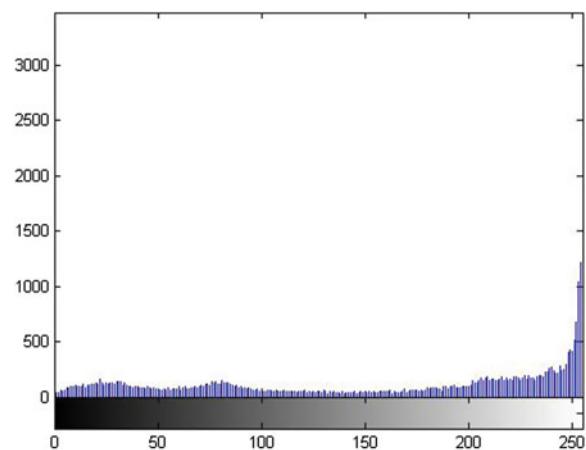
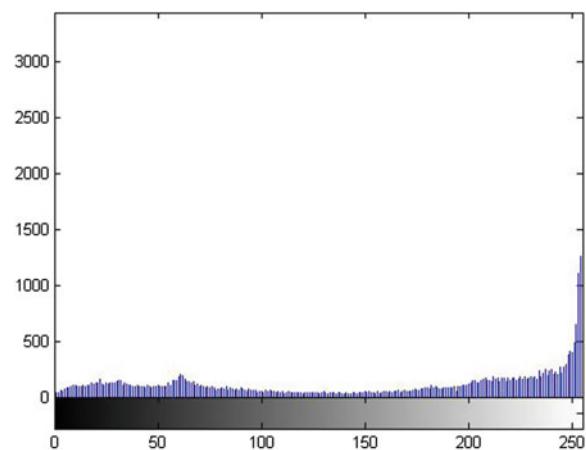
Fig. 1 Byke.jpg**Fig. 2** Red plane histogram**Fig. 3** Green plane histogram

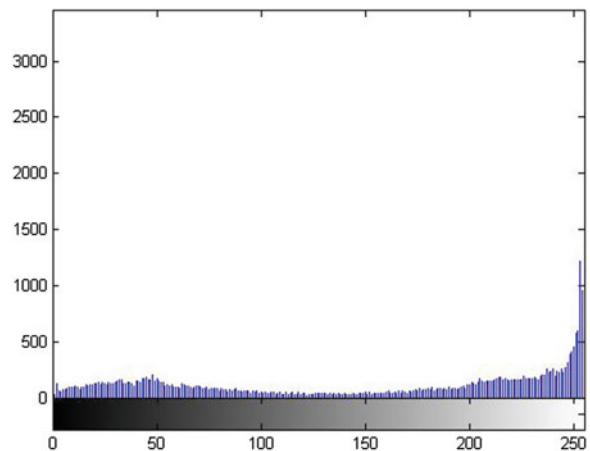
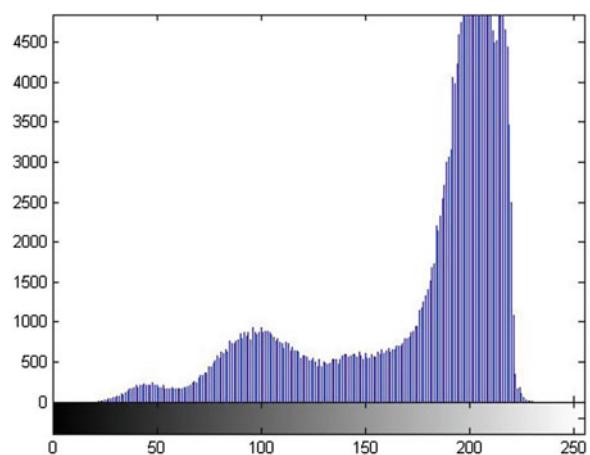
Fig. 4 Blue plane histogram**Fig. 5** Airplane.jpg**Fig. 6** Red plane histogram

Fig. 7 Green plane histogram

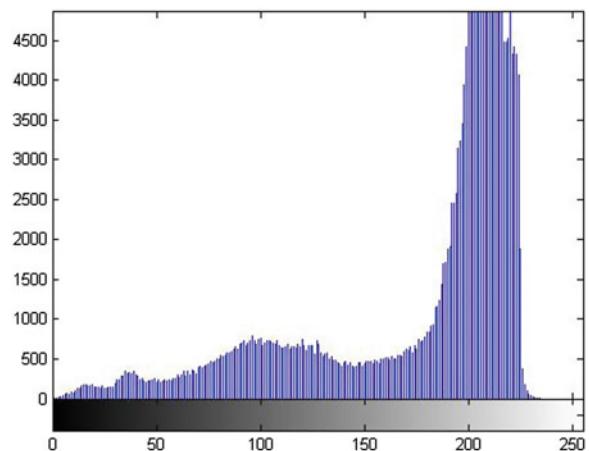


Fig. 8 Blue plane histogram

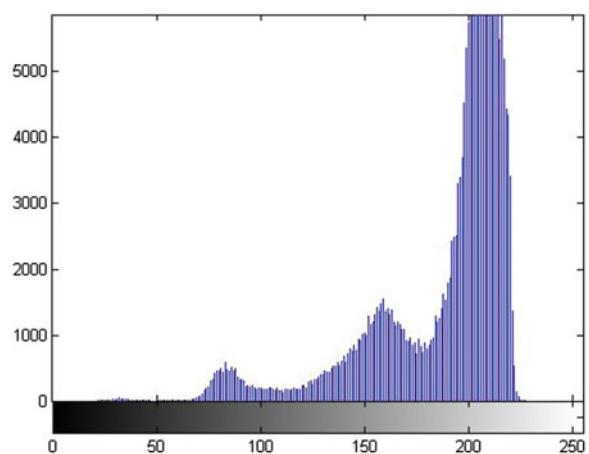


Fig. 9 Pepper.jpg

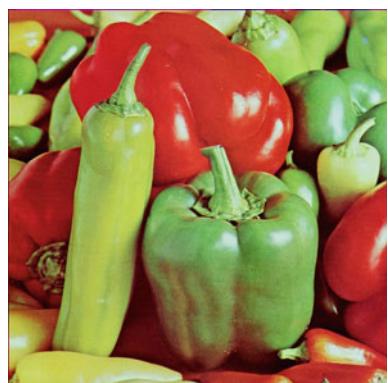


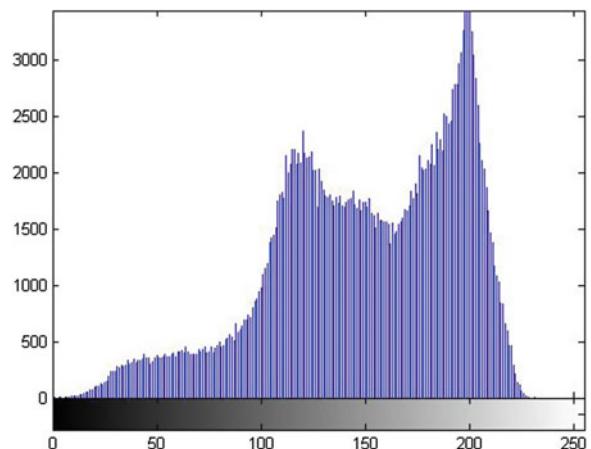
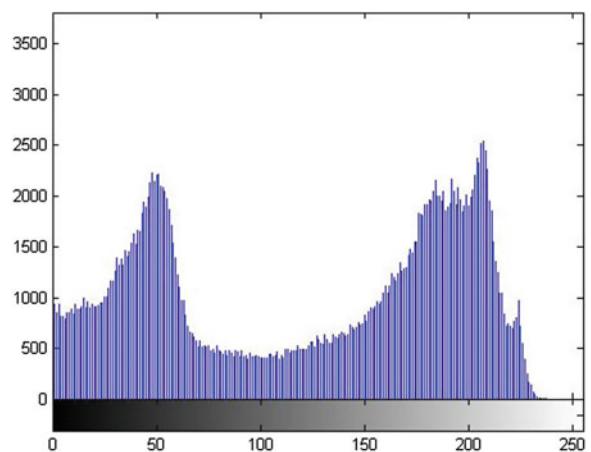
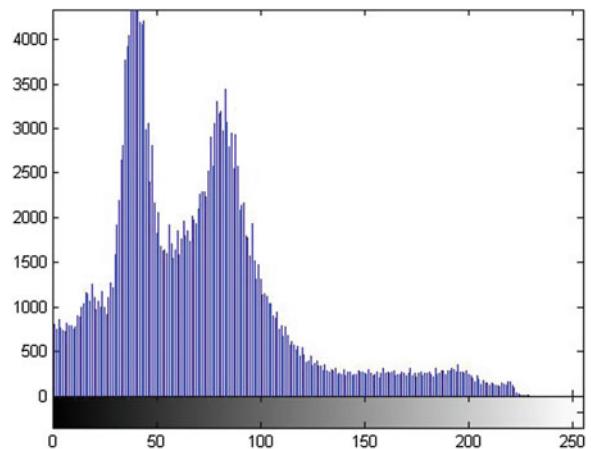
Fig. 10 Red plane histogram**Fig. 11** Green plane histogram**Fig. 12** Blue plane histogram

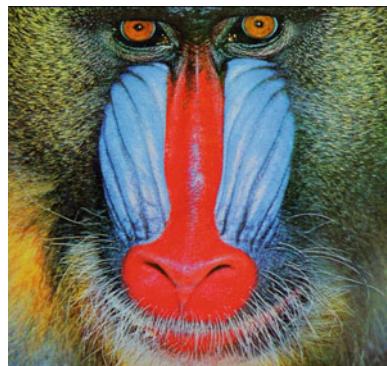
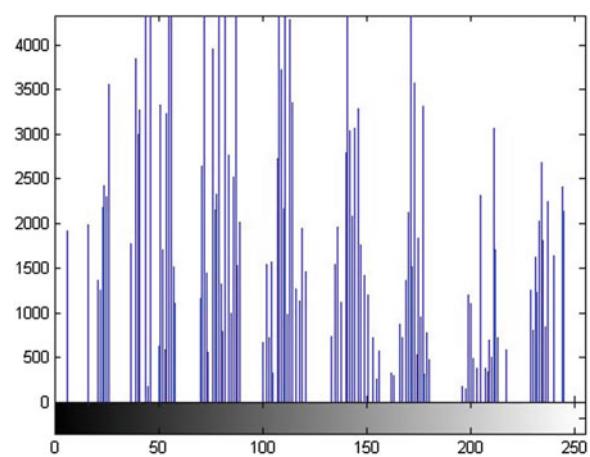
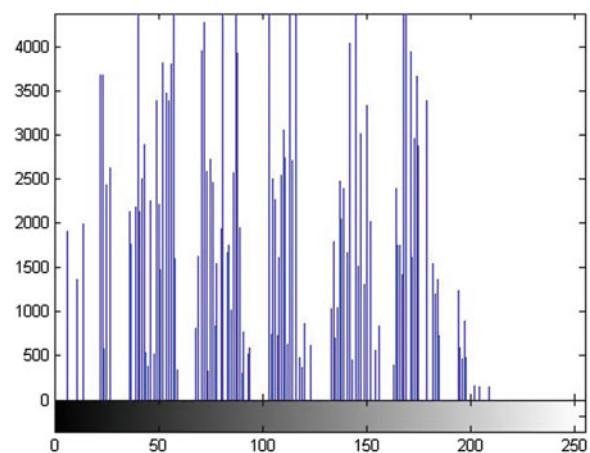
Fig. 13 Doctors payment ND**Fig. 14** Red plane histogram**Fig. 15** Green plane histogram

Fig. 16 Blue plane histogram

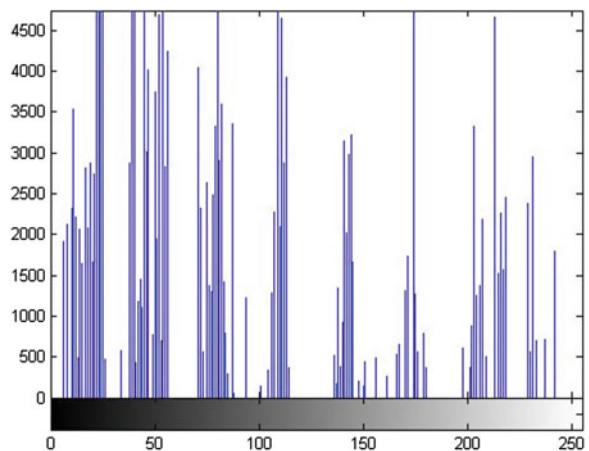


Fig. 17 Segmented image of byke.jpg



Fig. 18 Segmented image of airplane.jpg



Fig. 19 Segmented image of pepper.jpg



Fig. 20 Segmented image of babbon.jpg



6 Conclusions and Future Directions

In this paper, we have presented a noble approach for optimization-based color image segmentation. Here, we have used Kapur's Entropy Function as fitness function. Visual quality of the images shows that quality of segmentation is quite good. Now, this work can be extended into two different directions. Firstly, the algorithm (i.e., Cauchy-Mutated PSO) which is used for heuristic search that can be modified to propose new algorithms which may lead to better threshold value that yields better segmentation. In another direction, the fitness function (here Kapur's Entropy Function) can be changed by introducing new fitness functions such as Quadratic Entropy Function or PAI & PAI Entropy Function. Due to page limitation, we cannot provide histograms for images in three different color planes. Even new entropy functions can be defined which can be used as a fitness function and satisfies properties of basic entropy.

References

1. Gonzalez, R.C., Woods, R.E.: Digital image processing. Pearson Publication, USA (2007)
2. Pratt, W.K.: Digital image processing. Wiley, New York (2007)
3. Datta, R., Li, J., Wang, J.Z.: Content-based image retrieval: approaches and trends of the new age. In: Proceedings of the 7th ACM SIGMM International Workshop on Multimedia Information Retrieval, ACM, New York (2005)
4. Pham, D.L., Chenyang, X., Prince, J.L.: Current methods in medical image segmentation 1. *Ann. Rev. Biomed. Eng.* **2**(1), 315–337 (2000)
5. Miura, N., Nagasaka, A., Miyatake, T.: Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification. *Mach. Vis. Appl.* **15**, 194–203 (2004)
6. Shahin, A., Moudani, W., Chakik, F.: Image compression approach using segmentation and total variation regularization. In: Proceedings of the 2013 International Conference on Electronics, Signal Processing and Communication Systems (2013)
7. Ukinkar, V.G., Samvatsar, M.: Object detection in dynamic background using image segmentation: a review. *Int. J. Eng. Res. Appl. (IJERA)*, ISSN 2248-9622, **2**(3), 232–236 (2012)
8. He, Z., et al.: Toward accurate and fast iris segmentation for iris biometrics. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(9), 1670–1684 (2009)
9. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: IEEE International Conference on Neural Networks (ICNN95), IEEE Press, Australia, pp. 1942–1948 (1995)
10. Kapur, J.N., Sahoo, P.K., Wong, A.K.C.: A new method for gray-level picture thresholding using the entropy of the histogram. *J. Comput. Vis. Graph. Image Process.* **29**(3), 273–285 (1985)
11. Misra, D., Bose, I., De Chandra, U., Pradhan, B.: A multilevel image thresholding using particle swarm optimization. *Int. J. Eng. Tech.* **6**(2), 698–705 (2014)
12. Zhang, Q., Li, C., Liu, Y., Kang, L.: Fast multi-swarm optimization with cauchy mutation and crossover operation. In: Advances in Computation and Intelligence, pp. 344–352. Springer, Berlin (2007)

Large-Scale Global Optimization Using Dynamic Population-Based DE

Seema Chauhan, Suman Banerjee and Nanda Dulal Jana

Abstract Large-scale global optimization is one of the most challenging problems in the domain of stochastic optimization. Due to high dimensionality in the entire optimization process, different types of problems may occur for finding the global optima, e.g., solution space increases exponentially, problem complexity increases, and candidate search direction also increases exponentially. So, deterministic optimization algorithms cannot perform well for this kind of problems. Differential evolutionary algorithm is a population-based, stochastic search and optimization algorithm which can be used for global optimization problems. In this paper, we present self-adaptive dynamic population-based differential evolutionary algorithm which automatically adapts its parameters including population size.

Keywords Differential evolutionary algorithm · Large-scale global optimization · Amplification factor · Mutation probability · Crossover probability

1 Introduction

Large-scale global optimization [1, 2] is a kind of optimization problem over continuous search space having large dimensions. Our goal is to find out a D -dimensional vector $(x_1, x_2, x_3, \dots, x_i, \dots, x_D)$ which gives minimum value for the function. Particle swarm optimization (PSO) [3], simulated annealing, and some other global search and optimization algorithms perform well when problem dimension is small (like $D = 20, 30$). But in case of large problem (that contains large number of decision variables),

S. Chauhan (✉) · S. Banerjee · N.D. Jana

Department of Information Technology, National Institute of Technology,
Durgapur, West Bengal, India
e-mail: seemachauhan042@gmail.com

S. Banerjee
e-mail: banerjeesuman1991@gmail.com

N.D. Jana
e-mail: nandadulal.jana@it.nitdgp.ac.in

the size of the solution space is also increased exponentially as the number of decision variables increases, i.e., why these algorithms cannot perform well in case of large problem. Differential evolution is a simple and powerful technique to solve various optimization problems in continuous search spaces. The simple DE is good for low-dimensional problems (e.g., $D \leq 100$), which is not appropriate for our real-world problems. So to solve the problems on a large scale, a different kind of DE is introduced, which uses floating-point encoding evolutionary algorithm for optimization.

In simple DE, there are three control parameters, i.e., mutation probability (P_m), crossover probability (P_c), and selection process, which are fixed during the entire optimization process. In SADE, crossover probability and amplification factor (F) are changed dynamically. And in the proposed algorithm, we do the reduction of population size dynamically. So, in the proposed algorithm, there are three control parameters, i.e., crossover probability (P_c), mutation probability (P_m), and population size (N_POP).

The remaining portion of the paper is organized as follows: Sect. 2 presents the basics of DE algorithm, Sect. 3 presents our proposed dynamic population-based DE algorithm, Sect. 4 presents experimental studies which include benchmark functions and parameter settings, Sect. 5 describes simulation and result of our proposed algorithm, and at last, Sect. 6 draws conclusion and gives future direction.

2 DE Algorithm

DE is a stochastic, population-based, nature-inspired optimization algorithm. It is introduced by Storn and Price in 1996 [4, 5]. It is developed to optimize real parameter value function. The objective function is $f : X \subset R^D \Rightarrow R$ where the feasible region X is not empty, i.e., $(X \neq \emptyset)$. The minimization problem is to find $x^* \in X$ such that $f(x^*) \leq f(x) \quad \forall x \in X$.

In many real-world problems where the objective functions are non-differentiable, non-continuous, nonlinear, noisy, flat, multidimensional or have many local minima is difficult to solve analytically, i.e., why we may use DE to find the most optimal solution of this kind of problem. In comparison by Storn and Price in 1997, DE was more efficient than simulated annealing and genetic algorithms. After this, in 2004, Ali and Torn found that DE was both more accurate and more efficient than other genetic algorithms. In 2004, Lampinen and Storn demonstrated that DE was more accurate than many other optimization methods including genetic algorithm, simulated annealing, and evolutionary programming. The recent application of DE is in optimization of strategies for checkers. It is used in digital filter [6] and also in optimization of fermentation of alcohol.

3 Proposed Algorithm

Brest et al. [7] had proposed large-scale global optimization using self-adaptive differential evolution (SADE) algorithm. They introduced the algorithm with dynamic

amplification factor (F) and crossover rate (CR). As these parameters are fixed in earlier versions of DE, so the performance is not that much effective. As they modified the parameters F , CR, i.e., why the algorithm perform better than previous versions of DE, i.e., DECC-G, DECC-G*, MLGC etc. In any evolutionary algorithm, finding the global optimum value depends upon the balance of exploration and exploitation. At the initial stages of optimization process, exploration is more desirable than exploitation. So at the initial stages, population size should be large enough. As generation progresses, exploitation is more desirable than exploration, so in that case, population size can be reduced, i.e., why in our algorithm we have reduced population size dynamically over generation. In the proposed algorithm, we are modifying self-adaptive DE with dynamic population size. The proposed algorithm itself select after how many generation it may reduce population. In earlier versions, it is fixed through p_{\max} , but we may reduce population by half in each iteration. Formally, our algorithm is presented in Algorithm 1.

Algorithm 1: DE with Dynamic pop size

```

Input: Initialize current solution ( $X^c$ ) with dimension D and some constants  $F_l$ ,  $F_u$ ,  $T_1$ ,  $T_2$ 
Output: Minimum Functional Value.
begin
    for  $i = 1$  to  $N\_POP$  do
         $\downarrow$   $Fitness(i) = func(X(i,:), func\_no)$ 
    for  $i = 2$  to  $G$  do
        for  $i = 1$  to  $N\_POP$  do
            if  $rand(0,1) < T_1$  then
                 $\downarrow$   $F(i) = F_l + rand(0, 1) * F_u$ 
            for  $i = 1$  to  $N\_POP$  do
                if  $rand(0,1) < T_2$  then
                     $\downarrow$   $CR(i) = rand(0,1)$ 
                for  $i = 1$  to  $N\_POP$  do
                    for  $i = 1$  to  $D\_POP$  do
                        perform Mutation
                        Crossover
                        if  $rand \leq CR(i) OR P == D\_POP$  then
                             $\downarrow$   $U(i,j) = V(i,j)$ 
                        else
                             $\downarrow$   $U(i,j) = X(i,j)$ 
                    Selection
                    Based upon selection we may choose best optimal value
                Dynamic generation setting
                 $index = find(temp\_fitness == median(temp\_array))$ 
            for  $i = 1$  to  $N\_POP$  do
                if  $i == index$  then
                     $\downarrow$  temp_raw=temp_fitness
                for  $i = 1$  to  $N\_POP$  do
                    for  $j = 1$  to  $D\_POP$  do
                        temp_new = temp_new +  $(X(i,j) - temp\_raw)^2$ 
                         $\sigma = 1/N\_POP * temp\_new$ 
                        if  $G = 200 OR G = 400 OR G = 600 OR G = 800$  then
                            for  $i = 1$  to  $N\_POP/2$  do
                                if  $Fitness(i) > Fitness(N\_POP/2)$  then
                                    temp_array =  $X(NP/2)$ 
                                    temp_var =  $Fitness(N\_POP/2 + i)$ 
                                    Fitness(i) =  $Fitness(N\_POP/2 + i)$ 
                                    Fitness( $N\_POP + i$ ) = temp_ver
    end

```

4 Experimental Studies

4.1 Benchmark Functions

We have applied our proposed algorithm on benchmark functions for the CEC2010 Special Session and Competition on Large-Scale Global Optimization [8]. We have considered benchmark functions $f_2, f_3, f_4, f_5, f_6, f_9, f_{10}, f_{16}$ for our proposed algorithm, and the result is shown in Table 1. The functions are scalable, and dimension of these functions is $D = 1,000$. Each benchmark function required 25 test runs. The function is described in Table 1. The function may be separable and non-separable, depending upon the problem complexity level. As we know, separable problems are easy to solve, while the non-separable ones are complex, so it is difficult to solve. So intelligent search algorithms are required to solve these types of problems.

4.2 Parameter Settings

In simple DE algorithm, these are the following control parameters, i.e., mutation probability (P_m), crossover probability (P_c), and amplification factor (F). Values of these parameters have much impact on the performance of this algorithm [9]. But in the proposed algorithm, initial values of these parameters are given, but with the progress of generation, the proposed algorithm automatically updates mutation probability (P_m), amplification factor (F), and population size (N_POP) [10]. We use the following parameter settings: $G = 1,200$, $N_POP = 100$, $T_1 = 0.1$, $T_2 = 0.1$, $F_1 = 0.1$, $F_u = 0.9$, $D = 1,000$, $P_{\max} = 4$. The proposed algorithm is implemented using MATLAB 7.11.0 (R2010b) which runs on Intel (R) Core (TM) 2 Duo CPU E8400 @ 3.00 GHz with 2-GB RAM on Windows 7 Home Premium platform.

5 Results and Discussion

Table 2 presents the best, mean, and worst evaluation metrics and standard deviation of the benchmark functions, respectively, using the four different algorithms, e.g., differential evolutionary algorithm (SIMPDE) [4], SADE algorithm, PSO algorithm, and dynamic population-based SADE (DPOPDE) over 25 independent runs. The best results are marked in boldface. In case of the functions f_2, f_3, f_6, f_{10} and f_{16} , the proposed algorithm performs better and all these functions bare same nature of multimodality and shiftedness. For rest of the functions, performance of the proposed algorithm is quite satisfactory and difference of optimal values with other algorithms is very small. Convergence graphs for the proposed algorithm are also provided in the last for the benchmark functions $f_2, f_3, f_4, f_5, f_6, f_9, f_{10}, f_{16}$.

Table 1 The benchmark functions

Function	Mathematical representation	Range	Option
Shifted Rastrigin's fun	$F_2(x) = F_{\text{rastrigin}}(z) = \sum_{i=1}^D [z_i^2 - 10 \cos(2\pi z_i) + 10]$	[-5, 5]	0
Shifted Ackley's fun	$F_3(x) = F_{\text{ackley}}(z) = -20 \exp(-0.2 \sqrt{1/D \sum_{i=1}^D z_i^2}) - \exp(1/D \sum_{i=1}^D \cos(2\pi z_i)) + 20 - e$	[-100, 100]	0
m-rotated elliptic fun	$F_4(x) = F_{\text{rotelliptic}}[z(P_1 : P_m)] * 10^6 + F_{\text{elliptic}}[z(P_{m+1} : P_D)]$	[-100, 100]	0
m-rotated Rastrigin's fun	$F_5(x) = F_{\text{rotrastrigin}}[z(P_1 : P_m)] * 10^6 + F_{\text{rastrigin}}[z(P_{m+1} : P_D)]$	[-5, 5]	0
m-rotated Ackley's fun	$F_6(x) = F_{\text{rotackley}}[z(P_1 : P_m)] * 10^6 + F_{\text{ackley}}[z(P_{m+1} : P_D)]$	[-32, 32]	0
$D/2$ m group shifted and m-rotated elliptic fun	$F_9(x) = \sum_{k=1}^{D/2m} F_{\text{rotelliptic}}[z(P_{(k-1)*m+1 : P_{k*m}})] + F_{\text{elliptic}}[z(P_{D/2+1 : P_D})]$	[-100, 100]	0
$D/2$ m group shifted and m-rotated Rastrigin's fun	$F_{10}(x) = \sum_{k=1}^{D/2m} F_{\text{rotrastrigin}}[z(P_{(k-1)*m+1 : P_{k*m}})] + F_{\text{rastrigin}}[z(P_{D/2+1 : P_D})]$	[-5, 5]	0
D/m group shifted and m-rotated Ackley's Fun	$F_{16}(x) = \sum_{k=1}^{D/m} F_{\text{rotackley}}[z(P_{(k-1)*m+1 : P_k * m})]$	[-32, 32]	0

Table 2 Result of 8 functions

Function no.	Evaluation metrics	Algorithm			
		SIMPDE	DPOPDE	PSO	SADE
f_2	Best	1.83E+04	1.29E+04	1.88E+04	1.45E+04
	Worst	2.18E+04	1.68E+04	2.10E+04	1.68E+04
	Mean	1.99E+04	1.53E+04	1.99E+04	1.56E+04
	Std	8.77E+02	8.07E+02	6.48E+02	5.59E+02
f_3	Best	2.15E+01	2.05E+01	2.17E+01	2.07E+01
	Worst	2.15E+01	2.07E+01	2.17E+01	2.09E+01
	Mean	2.15E+01	2.06E+01	8.68E+01	2.08E+01
	Std	1.31E-02	6.35E-02	4.34E+00	4.39E-02
f_4	Best	2.25E+13	1.08E+13	3.44E+18	7.29E+12
	Worst	7.51E+13	4.16E+13	3.44E+18	5.08E+13
	Mean	4.44E+13	2.37E+13	1.38E+17	1.96E+13
	Std	1.23E+13	7.59E+12	6.88E+17	9.32E+12
f_5	Best	1.53E+08	1.55E+08	5.00E+13	1.16E+08
	Worst	5.24E+08	3.81E+08	5.00E+13	2.91E+08
	Mean	3.99E+08	2.46E+08	2.00E+12	1.99E+08
	Std	9.04E+07	6.49E+07	1.00E+13	4.90E+07
f_6	Best	2.77E+06	4.59E+06	2.28E+07	5.58E+06
	Worst	2.12E+07	2.11E+07	2.18E+07	2.11E+07
	Mean	1.68E+07	1.48E+07	4.36E+07	1.78E+07
	Std	7.03E+06	6.74E+06	9.75E+06	5.61E+06
f_9	Best	2.80E+010	1.18E+10	7.22E+13	7.54E+09
	Worst	7.32E+010	2.28E+10	7.22E+13	1.95E+10
	Mean	4.97E+010	1.66E+10	2.89E+12	1.10E+10
	Std	1.11E+010	3.15E+09	1.44E+13	3.15E+09
f_{10}	Best	1.95E+04	1.41E+04	9.99E+08	1.41E+04
	Worst	2.30E+04	2.72E+04	9.99E+08	1.81E+04
	Mean	2.09E+04	1.63E+04	9.99E+07	1.63E+04
	Std	8.02E+02	2.32E+03	3.16E+08	8.63E+02
f_{16}	Best	4.22E+02	4.20E+02	4.29E+02	4.20E+02
	Worst	4.31E+02	4.30E+02	4.30E+02	4.30E+02
	Mean	4.29E+02	4.23E+02	4.29E+02	4.23E+02
	Std	2.09E+00	2.70E+00	2.85E+00	2.65E+00

6 Conclusions and Future Directions

The proposed algorithm shows better performance in benchmark functions having multimodal landscape and shifted in nature. Now, this work can be extended in two different directions. Firstly, DE algorithm can be modified in some other way which may lead to good result like one can apply cooperative coevolutionary approach to DE. In another way, new parameter settings for this algorithm may also lead to improved results. The algorithm depends upon the control parameters (F , CR, and N_POP) during the optimization process. The result of the proposed algorithm in this paper gives the evidence that DPOPDE is a competitive for high-dimensional problem (Figs. 1, 2, 3, 4, 5, 6, 7, and 8).

Fig. 1 Convergence graph for F2

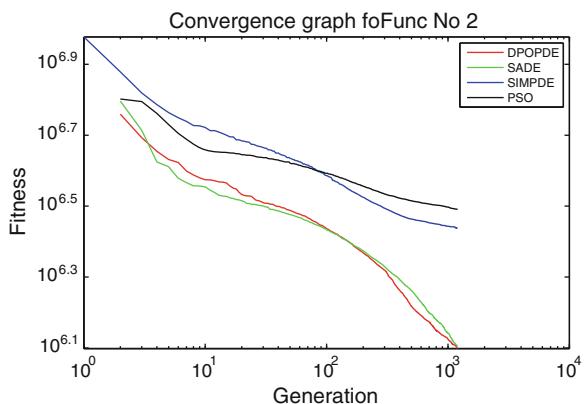


Fig. 2 Convergence graph for F3

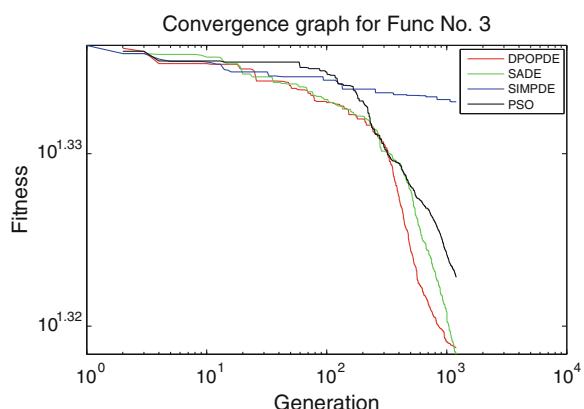


Fig. 3 Convergence graph for F4

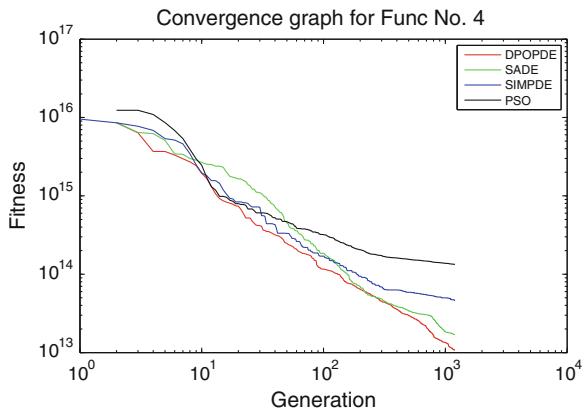


Fig. 4 Convergence graph for F5

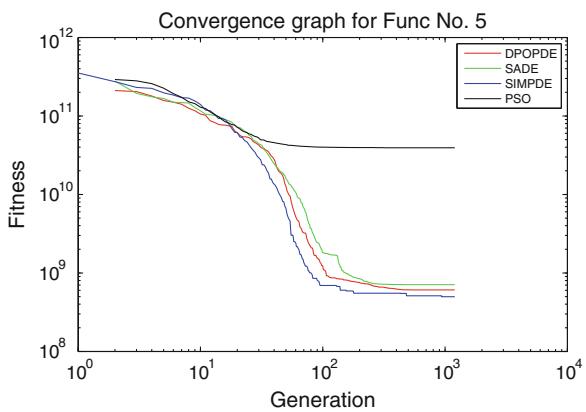


Fig. 5 Convergence graph for F6

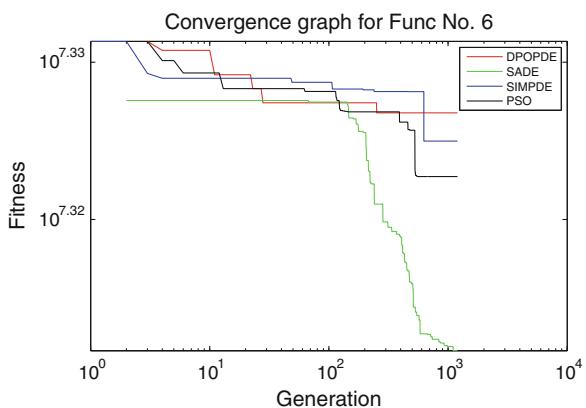


Fig. 6 Convergence graph for F9

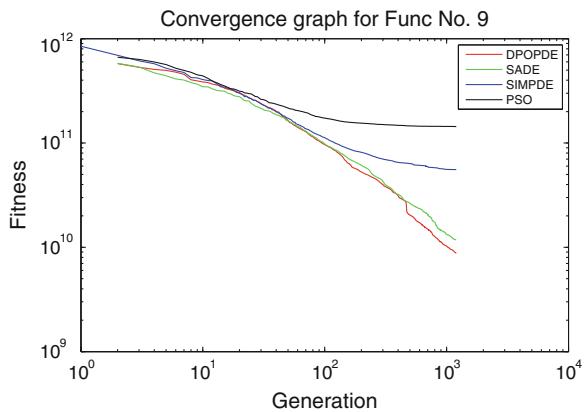


Fig. 7 Convergence graph for F10

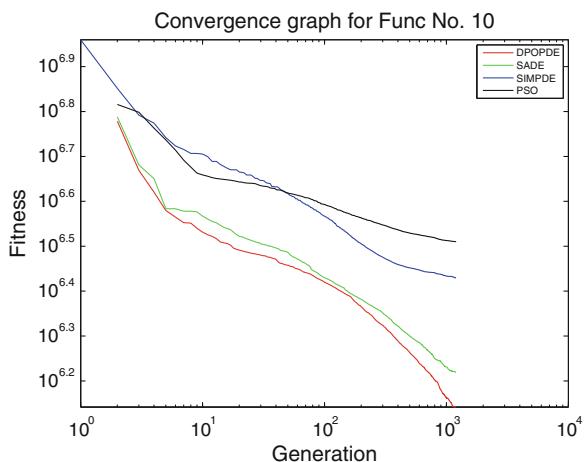
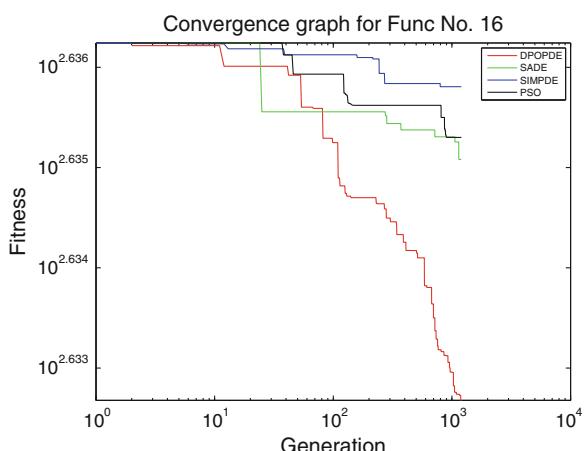


Fig. 8 Convergence graph for F16



References

1. Zamuda, A., Brest, J., Boskovic, B., Zumer, V.: Large scale global optimization using differential evolution with self-adaptation and cooperative co-evolution. In: IEEE Congress on Evolutionary Computation (2008)
2. Yang, Z., Ke, T., Yao, X.: Differential evolution for high-dimensional function optimization. In: IEEE Congress on Evolutionary Computation, pp. 3523–3530 (2007)
3. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: IEEE International Conference on Neural Networks (ICNN'95), pp. 1942–1948. IEEE Press, Australia (1995)
4. Storn, R., Price, K.: Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optim.* **11**, 341–359 (1997)
5. Das, S., Abraham, A., Konar, A.: Particle swarm optimization and differential evolution algorithms: technical analysis, applications and hybridization perspectives. In: Advances of Computational Intelligence in Industrial Systems, pp. 1–38. Springer, Berlin (2008)
6. Luitel, B., Venayagamoorthy, G.K.: Differential evolution particle swarm optimization for digital filter design. In: IEEE Congress on Evolutionary Computation (2008)
7. Brest, J., Zamuda, A., Fister I., Maucec, M.S.: Large scale global optimization using self-adaptive differential evolution algorithm. In: IEEE Congress on Evolutionary Computation (2010)
8. Tang, K., Li, X., Suganthan, P.N., Yang, Z., Weise, T.: Benchmark Functions for the CEC2010 Special Session and Competition on Large-Scale Global Optimization
9. Pedersen, M.E.H.: Good Parameters for Differential Evolution (2010)
10. Huang, F., Wang, L., Liu, B.: Improved differential evolution with dynamic population size. In: Intelligent Computing, pp. 725–730. Springer, Berlin (2006)

A Novel Preprocessing Approach for Human Face Recognition Invariant to Illumination

U.K. Jaliya and J.M. Rathod

Abstract Human Face recognition is one of the widely used biometric techniques for face identification and verification. It includes several subproblems like illumination variation, expression changes, aging, occlusion, and rotation of face images. Varying illumination is one of the well-known and challenging problems in human face recognition applications. In this paper, we proposed a novel approach to solve varying illumination problems in face images. The different stages include adaptive histogram equalization (AHE), Gaussian filtering, Log transform, difference of AHE+Gaussian filtering+Log image, and AHE+Log image, and then, we perform normalization. We are using principle component analysis (PCA) method for face recognition. The experimental results of proposed approach are compared with existing approaches, and it shows that our approach improves the performance of recognition under varying illumination conditions on Yale Face Database B.

Keywords Adaptive histogram equalization (AHE) · Principle component analysis (PCA) · Yale face database B · Gaussian filter · Log filter

1 Introduction

A biometric recognition system is an automated system that verifies or identifies a person's identity using a person's physiological characteristics and/or behavioral characteristics [1]. Face recognition has been growing rapidly in the past few years for its multiple uses in the areas of law enforcement, biometrics, security, and other

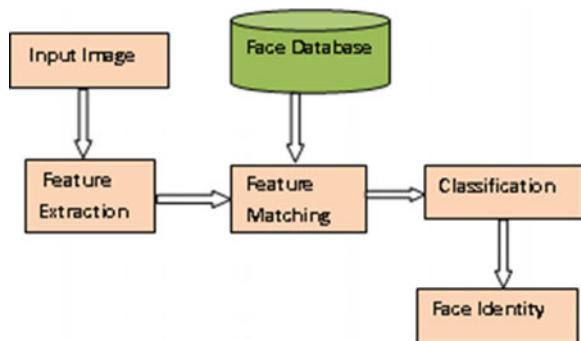
U.K. Jaliya (✉)

Research scholar, School of Technology, RK University, Rajkot, Gujarat, India
e-mail: udesang.jaliya@bvmengineering.ac.in

U.K. Jaliya · J.M. Rathod

BVM Engineering College, V.V.Nagar, Anand, Gujarat, India
e-mail: jmrathod@bvmengineering.ac.in

Fig. 1 Face recognition system [3]



commercial uses. As one of the most successful applications of image analysis and understanding, face recognition has recently gained significant attention, especially during the past several years. There are at least two reasons for such a trend: The first is the wide range of commercial and law enforcement applications, and the second is the availability of feasible technologies after several years of research. Face is one of the most common parts used by people to recognize each other [2].

The general block diagram for any face recognition system is given in Fig. 1. General steps in any face recognition system as depicted in Fig. 1 are discussed below:

The topic seems to be easy for a human, where limited memory can be a main problem, whereas the problems in machine recognition are manifold. Some of possible problems for a machine face recognition system are mainly [3]:

1. Facial expression change: A smiling face, a crying face, a face with closed eyes, and even a small nuance in the facial expression can affect facial recognition system significantly.
2. Illumination change: The direction where the individual in the image has been illuminated greatly affects face recognition success. A study on illumination effects on face recognition showed that lighting the face bottom up makes face recognition a hard task.
3. Aging: Images taken some time apart varying from 5 min to 5 years change the system accuracy seriously.
4. Rotation: Rotation of the individual's head clockwise or counter clockwise (even if the image stays frontal with respect to the camera) affects the performance of the system.
5. Size of the image: A test image of size 20×20 may be hard to classify if original class of the image was 100×100 .
6. Frontal versus profile: The angle in which the photo of the individual was taken with respect to the camera changes the system accuracy.

2 Novel Approach for Preprocessing

2.1 Introduction

In our novel approach for preprocessing, the input image is given as an input to the adaptive histogram equalization (AHE). It computes the histogram of a local image region centered at a given pixel to determine the mapped value for that pixel; this can achieve a local contrast enhancement. The output of the AHE is given as an input to the Gaussian filter. The Gaussian smoothing operator is a 2D convolution operator that is used to ‘blur’ images and remove detail and noise. It uses a different kernel that represents the shape of a Gaussian (‘bell-shaped’) hump.

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

where σ is the standard deviation of the distribution and $G(x, y)$ is the image. In our experiment, we have used $\sigma = 0.5$. The Gaussian is a high-pass filter, so it will remove low frequency and blur the image. The output image of AHE+Gaussian is given as an input to the Log transformation. Logarithmic transformations enhance low gray levels and compress the high ones. They are useful for non-uniform illumination distribution and shadowed images. Then, we are taking the difference of AHE+Gaussian+Log image and AHE+Log image, and finally, we perform normalization on the output image, so it will give the illumination-free image. The entire procedure is depicted in Fig. 2.

Our novel preprocessing approach output is given as an input to the face recognition system. Here, for experiment, we are using PCA method for face recognition and Euclidean distance as a classifier.

2.2 Face Database and Experimental Results

To illustrate the performance of our approach, we have conducted face recognition experiments using the Yale Face Database B. The database consists of images taken under 64 different lighting conditions, which are divided into 5 subsets according to

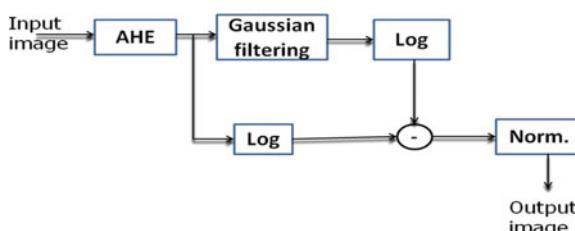


Fig. 2 Block diagram of a novel preprocessing approach



Fig. 3 Example of Yale B subset sample images

lighting angle. In Fig. 3, we show examples of face images in each subset. We used 640 images in total taken of 10 individuals in a frontal pose.

In our experiments, we use five subsets and each subset has different number of images with same lighting angle. These subset images are used for training set, and we are selecting few images within the same subset for test images. Figure 4 shows the output of each steps of Fig. 2, a novel approach block diagram.

From Fig. 4, we can see that the normalized image is the illumination-free image as compared to input image. Figure 5 shows the original images and output of our novel approach images. Row-1 shows original images, and row-2 shows processed images with our novel approach.

We have selected some of the images from the subsets as training set and testing set images. These images are used for face recognition using PCA method, and we



Fig. 4 Output of each steps of novel approach

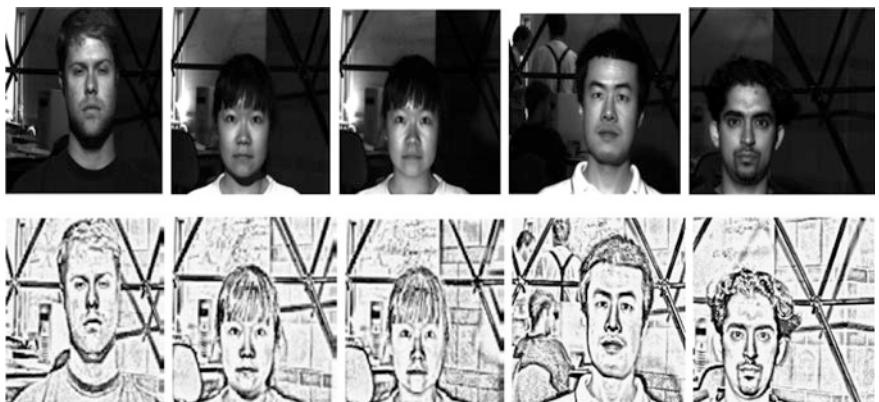


Fig. 5 Original images and novel approach output images

Table 1 Recognition rate

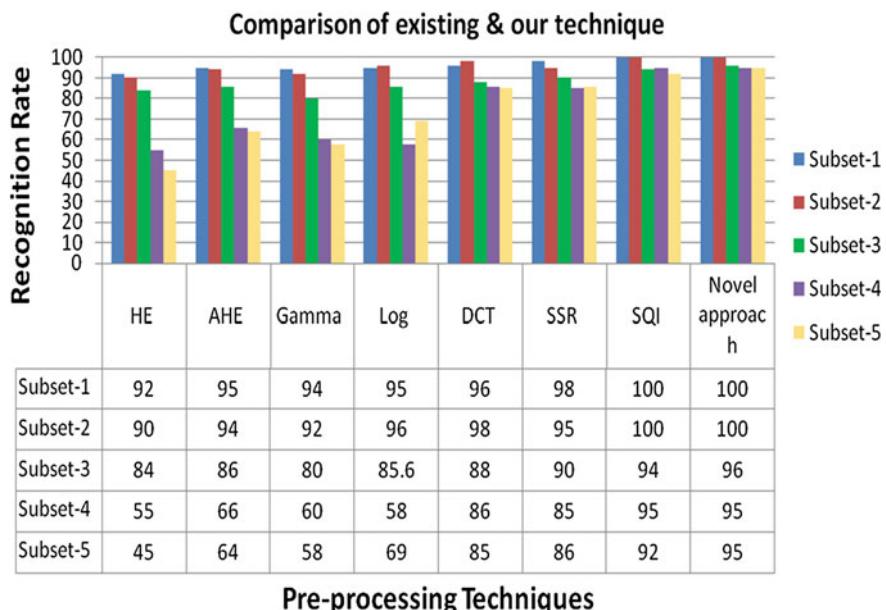
Subset	# of train images	# of test images	Recognition rate (%)
1	110	30	100
2	80	20	100
3	95	25	96
4	80	20	95
5	140	40	95

have calculated recognition rate as shown in Table 1: first column shows subset number, second shows number of training set images, third shows number of testing set images, and last column shows the recognition rate.

2.3 Comparison with Existing Methods [1–14]

Our novel preprocessing approach is compared with other approaches suggested by various others for face recognition using Yale Face Database B. Figure 6 shows the comparison of preprocessing approaches graphically. The graph in Fig. 6 indicates recognition rate in percentage.

From Fig. 6, we see that our method give good recognition rate compare to existing preprocessing techniques.

**Fig. 6** Comparison of various preprocessing approaches

3 Conclusion

In this paper, we presented a novel approach for preprocessing to illuminate the lighting effect from the face images. The experiment results show that our approach can significantly improve the recognition rates of face images under different lighting conditions compared with existing approaches. The recognition rate claimed by each of researcher depends on the database used and the number of subjects on which recognition task has been performed.

References

1. Anila, S., Devarajan, N.: Preprocessing technique for face recognition applications under varying illumination conditions. *Glob. J. Comput. Sci. Technol. Graph. Vis.* **12**(11), Version 1.0 (2012). Online ISSN 0975-4172 and Print ISSN 0975-4350
2. Jaliya, U.K., Rathod, J.M.: A survey on human face recognition invariant to illumination. *Int. J. Comput. Eng. Technol. (IJCET)* **4**(2), 517–525 (2013). Online ISSN 0976-6375 and Print ISSN 0976-6367
3. Jaliya, U.K., Rathod, J.M.: Illumination reduction techniques in human face recognition: a technical review. *Int. J. Graph. Image Process.* **3**(4) (2013)
4. Nam, G.P., Park, K.R.: New fuzzy-based retinex method for the illumination normalization of face recognition. *Int. J. Adv. Rob. Syst.* **9**(103) (2012)
5. Pizer, S.M., Amburn, E.P.: Adaptive histogram equalization and its variations. *Comput. Vis. Graph. Image Process* **39**, 355–368 (1987)
6. Zhichao, L., Joo, E.M., Nanyang Technological University Singapore: Face recognition under varying illumination. (2011)
7. Liu, H., Gao, W., Miao, J., Zhao, D., Deng, G., Li, J.: Illumination compensation and feedback of illumination feature in face detection. In: Proceedings of IEEE International Conferences on Info-tech and Info-net, vol. 23, pp. 444–449 (2001)
8. Xiong, Z.: An Introduction to Face Detection and Recognition. Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign (2009)
9. http://en.wikipedia.org/wiki/Difference_of_Gaussians
10. Turk, M., Pentland, A.: Eigenfaces for recognition. *J. Cogn. Neurosci.* **3**(1), 71–86 (1991)
11. Li Y., Wang, C., Ao, X.: Illumination Processing in Face Recognition. Shanghai Institute of Applied Physics, Chinese Academy of Sciences China (2010)
12. Gonzales, R.C., Woods, R.E.: Digital Image Processing, 2nd edn. Prentice Hall, Upper Saddle River (1992). ISBN-10: 0201180758
13. Shermina, J.: Illumination invariant face recognition using discrete cosine transform and principal component analysis. In: IEEE Proceedings of ICETECT (2011)
14. Yu, C., Jin, Z., Hao, C.: Illumination normalization based on different smoothing filters quotient image. In: IEEE Third International Conference on Intelligent Networks and Intelligent Systems, pp. 28–31 (2010)

Modeling and Simulation of Semi-active Jerk-Driven Damper on Severe Speed Breaker

C.K. Nirala and S. Mandal

Abstract In this study, performance of an analytical control strategy for semi-active suspension system called ‘jerk-driven damper’ (JDD) is analyzed. The analysis composes modeling of JDD control and MATLAB simulation considering a very realistic road input. Three continuous ‘positive amplitude half sinusoidal’ severe bumps are considered as input to the vehicle. For the ease of simulation, a semi-active damper is assumed to be a single degree of freedom (1DOF) incorporated into quarter-car model which is a subject of base excitation. The optimality in terms of vertical acceleration (comfort objective) of sprung mass for JDD is examined over ‘passive,’ ‘skyhook’ (SH), and ‘acceleration-driven-damper’ (ADD) control. The performance of JDD is found to be better on comparison.

Keywords Semi-active suspension system • Skyhook • Jerk-driven damper • Acceleration-driven-damper • Severe bumps

1 Introduction

Suspension system of a vehicle is a system of linkage, springs, and shock absorbers used to connect the vehicle chassis with the axle of the wheels. The primary purpose of any suspension system is to make the body of the vehicle isolated from the road irregularities which is considered to be basically a disturbance to the passengers sitting in vehicle. A passive suspension system containing a spring and a damper is the simplest way to isolate the vehicle from such disturbances. There is some trade-off between handling and ride quality in the passive suspension because

C.K. Nirala (✉)

Indian Institute of Technology Patna, Patna 800013, India
e-mail: nirala@iitp.ac.in

S. Mandal

Birla Institute of Technology Patna, Patna 800014, India
e-mail: soumenbhu@gmail.com

the system cannot fulfill the requirement of high damping value at the resonance and a low damping value at high frequency of input disturbances [1]. The limitations associated with passive suspension system could easily be understood by study of transmissibility for suspension systems. In order to fulfill the requirement of variable damping values at varying frequencies of road disturbances, a semi-active suspension system is introduced. A semi-active suspension system consists a variable damper which has capability to change the rate of energy dissipation depending upon the condition of motion. So, to meet the right compromise, the tuning of damper between its ‘Off’ and ‘On’ mode is required to get the appropriate damping value. According to Yao et al. [2], a ‘magneto rheological’ (MR) damper is a highly sophisticated controlled damper for semi-active suspension system. In the last decades more focus has been given to the study of semi-active suspension system due to its advantages over passive suspension system. Vertical acceleration is identified as the performance parameter of any suspension system of vehicle [3]. It has been reported by various researchers that the better performance could be achieved, if one can minimize the vertical acceleration of the sprung mass. A number of possible two-state ‘On–Off’ and ‘continuous controlled’ semi-active suspension systems have been broadly studied particularly in the last few decades, and a consistent research activities are still involved to improve the performance. A control algorithm called ‘balance logic’ was first introduced by Rakheja et al. and further developed by a group of researchers [4, 5]. Karnopp et al. [5] derived control logic for semi-active suspension system called ‘skyhook’ (SH), in which, the damper is assumed to be hooked to a fixed point in the sky. However, the ‘skyhook’ suspension was introduced before ‘balance logic’ but later it was fully realized by Yi and Song [6]. A new logic in semi-active suspension control called ‘groundhook’ was introduced by Novak and Valasek [7] which aims at reducing dynamic tire force, thus improving the handling and reducing road damage at the same time. This logic was reported to be most suitable for heavy vehicle. Similar to the skyhook, the ‘groundhook’ damper is assumed to be hooked to a fixed point, here which is ground instead of sky. Further, the skyhook and ‘groundhook’ logics are combined together which results in reducing both body acceleration and dynamic tire force [8].

In all the above-mentioned literatures, the semi-active damping force is reported as proportional to chassis velocity. Based on the above discussed control strategies in semi-active suspensions, a new nonlinear stroke dependent damper design is proposed, in which the damping rapidly changes based on a new logic. This non-linear damper leads to improved performance and riding comfort by lowering the acceleration level for the sprung mass [9]. A variable damper as well as variable stiffness of semi-active suspension system has been discussed and has shown the feasibility of it in improved performance [10]. Saiful et al. [11] developed a modified SH suspension system by incorporating a current generator model which shows an improved performance after simulation. One innovative acceleration-based control strategy named ADD proposed by Sergio et al. [12] and found optimal over SH control. The innovation of ADD was based on continuous control of SH suspension. Nirala et al. [13] found some further scope to improve the semi-active suspension

system of light vehicle by introducing a new control algorithm called ‘jerk-driven damper’ (JDD) control. The JDD was simulated considering a very simple smooth road bump. But, the betterment in any suspension control can only be ensured if its performance is analyzed over a realistic road disturbance.

The objective of this work was set to analyze the performance of a JDD control considering three continuous severe road bumps as the input to the vehicle. The betterment in the performance is assured on observation of the comparison results with other semi-active control strategies. The performance is measured in terms of vertical acceleration (comfort objective) of the sprung mass.

2 Quarter-Car Model

A general model of quarter car incorporating a 2DOF system is shown in Fig. 1. If the stiffness is considered high, the tires of the vehicle behave almost as a rigid body and the equivalent of this is also shown in the same figure. Sprung mass in the figure is the portion of the vehicle’s total mass that is supported above the suspension, whereas the unsprung mass is the mass of wheels and other components directly connected with wheels.

For this equivalent quarter-car model, the differential equation could be written as follows:

$$m\ddot{x} = -b_s(\dot{x} - \dot{u}) - k(x - u) \quad (1)$$

Equation (1) is a nonlinear equation which governs the system shown in Fig. 1. The nonlinearity of the equation is caused by the variable semi-active damper (b_s).

Where m is the sprung (body) mass of quarter car. k and b_s are the fixed stiffness and variable damping of the suspension system, respectively. $x(t)$ and $u(t)$ are the sprung mass and base excitation, respectively.

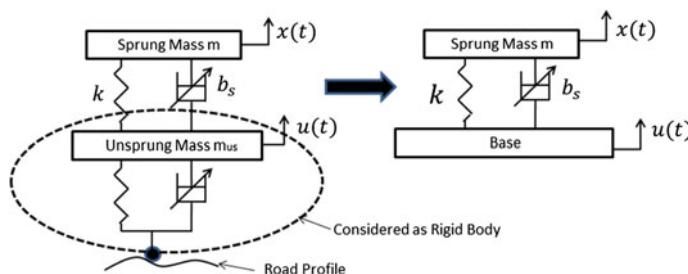


Fig. 1 A 2DOF quarter-car model and an equivalent 1DOF model [13]

3 Different Kinds of Existing Control Strategies

Above-mentioned simplified 1DOF quarter-car model is used to propose various kinds of semi-active control strategies. This section provides a brief of these existing semi-active suspension systems with their control algorithm.

Skyhook (SH) control: In ‘Skyhook’ suspension, there is an assumption of fixing the damper between chassis and sky. This is a widely studied suspension system which has already shown advantage over other semi-active suspensions such as passive, balance logic, and groundhook [4]. Figure 2a shows the virtual SH configuration of semi-active suspension system. The damper is assumed to be fixed in sky, and frequency response of this virtual system is shown in Fig. 2b. If we analyze the Fig. 2b, c, it could be observed that the frequency response of ‘passive suspension’ system at higher damping is good before cutoff point and very low damping is required after cutoff point. Frequency response plot of skyhook configuration shows the minimum deflection of sprung mass with all the damping at higher frequency. From this, it could be revealed that the requirement of low damping at high frequency can be achieved by making 1DOF ‘passive suspension’ system to its equivalent skyhook configuration. But, practical implementation of SH configuration is not possible. So, SH control algorithm gives same damping as SH configuration for passive 1DOF suspension system and shows its considerable better performance. The y -axis in Fig. 2b, c shows the transfer function of single degree of freedom system.

An on-off damper is basically a two-state damper which has a choice to shift the damping value either maximum or minimum. The maximum value of the damping is same as the damping value of passive damper. b_{min} and b_{max} refer to the maximum and minimum value of damping, and b_{sky} is damping in skyhook configuration which can be equal to the maximum value of damping. A SH algorithm approximation is an incorporation of two-state damper with a specific algorithm which is given in Eq. (2).

$$b_s(t) = \max \left[b_{min}, \min \left[b_{sky} \frac{\dot{x}}{(\dot{x} - \dot{u})} \right] \right], \text{ if } \dot{x}(\dot{x} - \dot{u}) \geq 0 \text{ otherwise } b_{min} \quad (2)$$

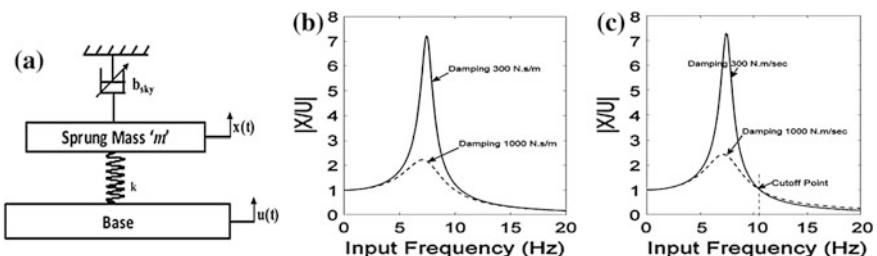


Fig. 2 a SH configuration of suspension system, b frequency response of equivalent SH suspension, and c frequency response of passive suspension (1DOF) system

ADD control: Similar concept of continuous controlled SH suspension can be implemented for the continuous control algorithm for the two-state on-off ADD suspensions and can be written as in Eq. (3).

$$b_s(t) = \max \left[b_{min}, \min \left[b_{sky} \frac{\dot{x}}{(\dot{x} - \ddot{u})} \right] \right], \quad \text{if } \ddot{x}(\ddot{x} - \ddot{u}) \geq 0 \text{ otherwise } b_{min} \quad (3)$$

ADD has shown a lower vertical acceleration than SH suspension [12]. ADD and skyhook have got a very similar model which is quite simple also. A comparative study of the two has been performed on the basis of MATLAB simulation, and the result is plotted in ‘simulation results and discussion section’.

4 JDD Control

Differentiating the Eq. 1 will give the third-order governing equation of 1DOF system in terms of jerk force which is shown in Eq. 4. To understand the JDD working principle more accurately, Eq. 4 is analyzed for the different possible cases of jerk force $m\ddot{x}$.

$$m\ddot{x} = \{-b_s(\ddot{x} - \ddot{u})\}_A \{-k(\dot{x} - \dot{u})\}_B \quad (4)$$

The equation could be divided into two parts A and B of algebraic terms given in above equation. The two possible cases are (1) When the jerk is greater zero and (2) when the jerk is negative.

In both the cases, the quantity A in Eq. 4 could be +ve or -ve depending upon the sign of the acceleration (\ddot{x}) in the equation. Now, in order to satisfy the equation and maintaining the minimum value of jerk force, there may be two possibilities in quantity B for each condition of quantity A . These two possibilities in quantity B lead in choosing the damper condition from its maximum and minimum value in both the cases. Figure 3 depicts the full explanations of different cases with damper switching decisions.

Now, let us observe the third-order deferential equation of the skyhook configuration.

$$m\ddot{x} = -b_{sky}(\ddot{x}) - k(\dot{x} - \dot{u}) \quad (5)$$

The fact of better frequency response in case of SH suspension is known now. Equating the Eqs. 4 and 5 will give the damping required for continuous control of damper shown in Eq. 6.

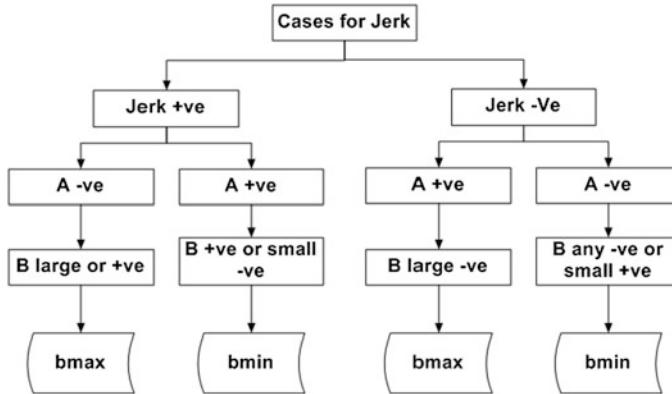


Fig. 3 Flow chart for damping value selection depending upon jerk force

$$b_s = b_{\text{sky}} \frac{\ddot{x}}{(\ddot{x} - \ddot{u})} \quad (6)$$

When the JDD is in ‘On’ condition, it gives the damping according to Eq. 6. But, we always require maximum damping at lower excitation frequency (refer to Fig. 2) and very low at higher frequency. Equation 6 will not be valid when $\ddot{x} = \ddot{u}$ and $\ddot{x} \gg \ddot{u}$, so for these cases, the damper has to be set maximum possible value of damping. The selection of maximum possible value could be done by using Eq. 7.

$$b_s = \max \left[b_{\min}, \min \left[b_{\text{sky}} \frac{\ddot{x}}{(\ddot{x} - \ddot{u})} \right] \right] \quad (7)$$

The fixed damping of passive suspension refers the maximum value of damping of it. After analyzing the cases discussed above, it is possible to propose a control strategy for the JDD suspension must be satisfying the cases. The control strategy can be seen in Eq. 8.

$$b_s(t) = \max \left[b_{\min}, \min \left[b_{\text{sky}} \frac{\ddot{x}}{(\ddot{x} - \ddot{u})} \right] \right], \quad \text{if } \ddot{x}(\ddot{x} - \ddot{u}) \geq 0 \text{ otherwise } b_{\min} \quad (8)$$

Table 1 is prepared for more clarification for both the cases.

The jerk force applied on the sprung mass is proportional to the relative acceleration, minimizing which is the comfort objective. MATLAB simulation is performed to get the time response of this vertical acceleration of sprung mass and compared with other well-established suspension strategies. The results have been discussed in section ‘simulation results and discussion’.

Table 1 Algorithm of the JDD control

Case	Jerk	Rel. acc.	Rel. velocity ($\dot{x} - \ddot{u}$)	Damper condition	Damping value
1	$m\ddot{x} \geq 0$	$(\ddot{x} - \ddot{u}) \geq 0$	-ve and larger in magnitude	On	b_s
		$(\ddot{x} - \ddot{u}) < 0$	Any -ve or small positive	Off	b_{\min}
2	$m\ddot{x} < 0$	$(\ddot{x} - \ddot{u}) < 0$	+ve and larger in magnitude	On	b_s
		$(\ddot{x} - \ddot{u}) \geq 0$	Any +ve or small negative	Off	b_{\min}

5 Simulation Results and Discussion

The JDD suspension discussed in previous section is now simulated in this section. MATLAB simulation is performed for time response of the sprung mass acceleration considering half sinusoidal severe road bump as input.

The following sets of parameters related to quarter car are used for simulation:

Sprung mass $m = 290$ kg, stiffness of the suspension $k = 16182$ N/m. $b_{\max} = 1,000$ N · s/m and $b_{\min} = 300$ N · s/m. where b_{\max} and b_{\min} are the maximum and minimum coefficients of on-off damper, respectively. The dimensions of the road bump is chosen as half sinusoidal with height, $H = 0.075$ m and width, $W = 0.4$ m each. The gap between two consecutive bumps were kept as $2W$ (Fig. 4).

The simulation has been done on a severe road breaker consists of three continuous bums as shown in Fig. 4. The comparison of JDD performance with passive, SH, and ADD suspension systems is shown in Fig. 5a. Table 2 shows a noticeable improvement in comfort objective (vertical acceleration of sprung mass) in case of JDD suspension. Although the performance is comparatively good, more settling (Fig. 5a) time of JDD suspension can be considered a disadvantage of this algorithm.

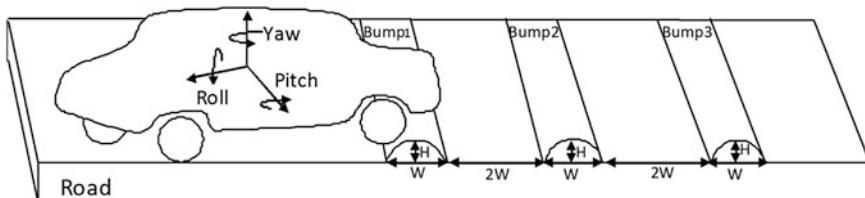


Fig. 4 A typical continuous road bump (severe) used for simulation

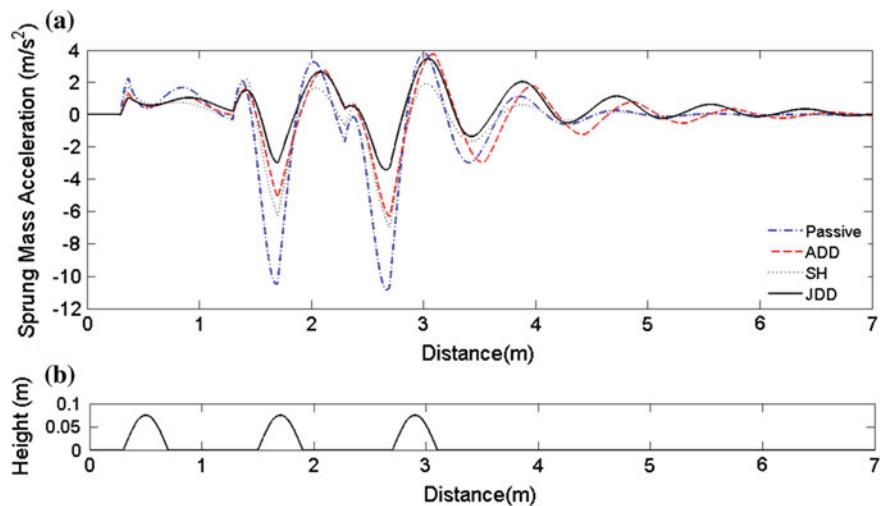


Fig. 5 a Time response of the sprung mass acceleration and b the severe breaker

Table 2 Sprung mass acceleration in various kind of semi-active suspensions

Suspension types	Passive	SH	ADD	JDD
Max. acc. magnitude (m/sec ²)	10.481	6.986	6.345	3.473

6 Conclusion

With the severe speed breaker (continuous road bumps) as an input to the vehicle, the JDD control algorithm of the suspension system shows comparatively better performance of sprung mass as a lower vertical acceleration. The sprung mass acceleration response in case of JDD is approximately 66.8 and 50.2 % less than the passive and SH suspension, respectively. The JDD is again has shown a better performance by 45.26 % improvement in vertical acceleration, when compared to ADD suspension. So, it can be concluded from the above discussion that JDD controlled semi-active suspension system performs better on severe speed breakers.

References

- Thomson, W.T., Marie, D.D., Chandramouli, P.: Theory of Vibration and Applications, vol. 5. CRC Press, Pearson Education, New York (1996)
- Yao, G.Z., Yap, F.F., Chen, G., Li, W.H., Yeo, S.H.: MR damper and its application for semi-active control of vehicle suspension system. Mechatronics **12**(7), 963–973 (2002)

3. Spelta, C., Previdi, F., Sergio, M.S., Bolzern, P., Cutini, M., Bisaglia, C., Simone, A.B.: Performance analysis of semi-active suspensions with control of variable damping and stiffness. *Veh. Syst. Dyn.* **49**(1–2), 237–256 (2011)
4. Rakheja, S., Sankar, S.: Vibration and shock isolation performance of a semi-active on-off damper. *Trans. ASME* **107**, 398–403 (1985)
5. Karnopp, D.C., Crosby, M.J., Harwood, R.A.: Vibration control using semi-active force Generators. *J. Manuf. Sci. Eng.* **96**(2), 619–626 (1974)
6. Yi, K., Song, B.S.: A new adaptive sky-hook control of vehicle semi-active suspensions. *Proc. IMechE J. Automob. Eng.* **213**, 293–303 (1999)
7. Novak, M., Valasek, M.: A new concept of semi-active control of truck's suspension. In: AVEC 96, Aachen, Germany (1996)
8. Guglielmino, E., Sireteanu, T., Charles, W.S., Ghita, G., Giuclea, M.: Semiactive Suspension Control. Springer material. ISBN 978-1-84800-230-2 (2008)
9. Raj, A.H.K., Padmanabhanr, C.: A new passive non-linear damper for automobiles. *Proc. IMechE Part D: Automob. Eng.* **223**, 1435–1443 (2009)
10. Spelta, C., Sergio, M.S., Fabbric, L.: Experimental analysis of a motorcycle semi-active rear suspension. *Control Eng. Pract.* **18**(11), 1239–1250 (2010)
11. Saiful, A., Abu, B., Jamaluddin, H., Rahman, A.R., Samin, P.M., Hudha, K.: Vehicle ride performance with semi-active suspension system using modified skyhook algorithm and current generator model'. *Int. J. Veh. Auton. Syst.* **6**(3–4), 197–221 (2008)
12. Sergio, M.S., Silani, E., Bittanti, S.: Acceleration-driven-damper (ADD): an optimal control algorithm for comfort-oriented semiactive suspensions. *J. Dyn. Syst. Meas. Contr.* **127**(2), 218–229 (2005)
13. Nirala, C.K., Kumar, R.P.: A new semi-active suspension system based on jerk driven damper control (JDD). In: Proceedings of ICCMM, IIT Guwahati, pp. 561–568 (2011)

Voltage-Controlled Ring Oscillator for Harmonic Frequency Generation

Surajit Mal, Ashis Kumar Mal and Sumalya Ghosh

Abstract In this paper, a voltage-controlled-based ring oscillator is proposed to generate different frequencies. Frequency of a ring oscillator depends on two factors: number of delay stages (length) in the ring and number of delay stages of the inverters, which can be controlled by a set of control voltages. In the proposed design of voltage-controlled ring oscillator, multiplexers are employed to control either length of the chain (N) or control voltage of the inverters or both to vary the delay hence output frequency. Three different types of pseudo NMOS/PMOS/NMOS with PMOS inverters-based ring oscillators are studied to generate harmonic such as frequencies. Simulation carried out in UMC 180 nm CMOS technology shows promising results.

Keywords Ring oscillator · Transmission gate · Frequency stability of ring oscillator

1 Introduction

Oscillatory behavior is being present everywhere, and this behavior has been employed in most clock generators circuits, frequency synthesizers for their quadrature or multiphase outputs, and easy integration in standard CMOS technologies. An oscillator provides a perfect time reference, but all physical oscillators are corrupted by undesired noise. A variety of oscillators are available, but frequency band of oscillation and performance of noisy environment are different from one class to another class of oscillators [1–3]. Voltage-controlled oscillator is

S. Mal (✉) · A.K. Mal · S. Ghosh
Department of Electronics and Communication Engineering,
National Institute of Technology, Durgapur 713209, India
e-mail: surajit3855@gmail.com

A.K. Mal
e-mail: a.k.mal@ieee.org

S. Ghosh
e-mail: sgnitd@yahoo.com

an electronic oscillator designed to control oscillation frequency by controlling voltage input. The basic building block in entire electronic industry is VCO-based oscillator [4]. The architecture of VCO falls under two categories: LC oscillator which is composed of active devices, coupled with LC resonant circuit and Ring oscillator which is composed of cascaded delay element with a positive feedback. LC tanks VCOs have superior phase-noise performance, but they require massive passive resonators and limited tuning range. Ring-based VCO is applicable owing to their ease of integration, small die area but due to poor phase-noise performance, they have marginal use in low phase-noise field such as wireless communication and optical transceivers [5, 6].

In the face of shrinking feature size the supply voltage reduced by several range, a conventional tuning scheme will result expensive voltage-to-frequency (K_{VCO}) gain and also increases the noise sensitivity [7]. VCO must be capable of transfer input voltage to oscillation frequency. An ideal VCO is a circuit whose output frequency depends on its control voltage (V_a).

$$\omega_{VCO} = \omega_0 + K_{VCO} \cdot V_a \quad (1)$$

ω_0 is the free running frequency without control voltage. V_a is the input of VCO that controls the desired frequency. K_{VCO} is the gain of VCO and is given by the following:

$$K_{VCO} \geq 2\pi \frac{\omega_2 - \omega_1}{V_2 - V_1} \quad (2)$$

To avoid such problem, various designs are approached to find the frequency range. However most Ring oscillators implement the fine tuning through controlling drain current (I_{SS}) of the circuit, drain current can be controlled by changing the bias voltage (V_a). To get multiple bias voltage digitally, multiplexer and transmission gate are proposed here [4].

This paper presents a design of voltage-controlled ring oscillator using transmission gates with various frequency patterns at low power supply 1.8 V. The frequency tuning is implemented by adjusting the current through the drain terminal. Different delay cell circuits are developed which are discussed in Sect. 3. This paper represents on harmonic frequency generation where frequency deviation is in uniform.

2 Ring Oscillator

A Ring oscillator is formed by using odd or even number of inverter or delay stages (N) which are couples in a positive feedback loop. To generate oscillation, if one node of the ring is excited, then the pulse propagates through the delay cell and after that reverse polarity is formed in initially excited node. The ring must provide a phase shift of 2π and has unity voltage gain [2]. Each delay stage must provide a phase shift of $\frac{\pi}{N}$ and remaining π phase shift is provided by DC inversion. A single-inverting stage

does not oscillate because it is open-loop circuit which contains single pole, thereby providing maximum phase shift of 90° and due to signal inversion from gate to drain a dc phase shift of 180° . The maximum phase shift is 270° ; due to these loops therefore it fails to sustain oscillation. At high frequency loop gain vanishes, so even stage does not satisfy Barkhausen's criteria at high frequency, so it's failing to oscillate [8, 9]. This paper presents a design of a CMOS ring VCO based on transmission gates with wide tuning range 100 MHz–1600 MHz at low power supply 1.8 V [10]. Three inverting stages are enough for sustained oscillation provided gain of each stage is sufficient. The basic building block of 5-stage single-ended pseudo ring oscillator is depicted in Fig. 1.

To determine the oscillation frequency, let the delay of each inverter is τ_d . Input signal must pass through N number of stages, so for first cycle, the output is high and total delay time is $N \cdot \tau_d$. Now for next cycle, the output is low and delay is same as previous. So, resulting clock cycle or total period is $2N\tau_d$. The oscillation frequency of excited node is given by following

$$f_{\text{osc}} = \frac{1}{2N\tau_d} = \frac{I_{\text{ss}}}{2NC_L V_{\text{SW}}} \quad (3)$$

From above equation, the unknown parameter is, τ_d , the propagation delay. The value of τ_d is estimated in next section.

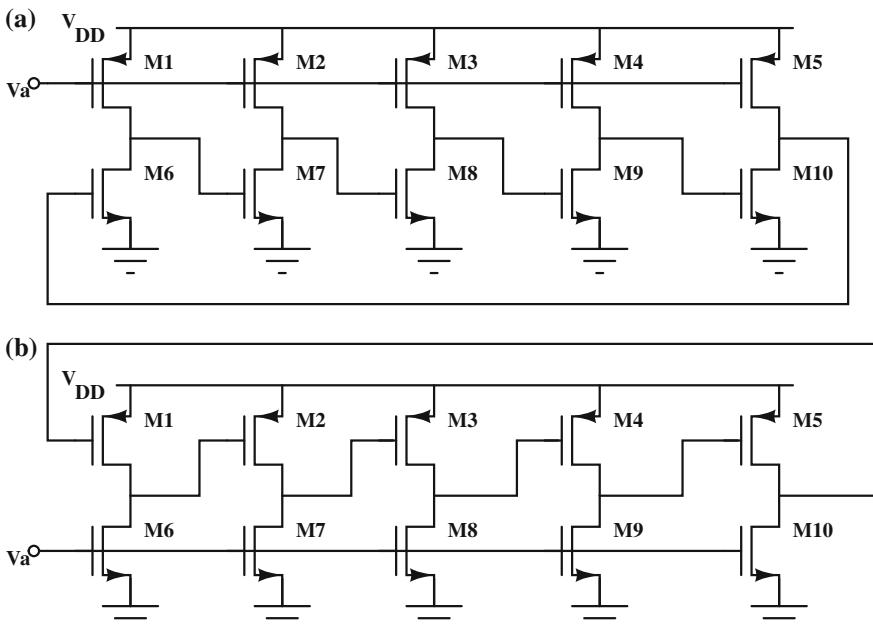


Fig. 1 Primitive delay cells for VCO-based ring-type design **a** Pseudo NMOS ring oscillator and **b** Pseudo PMOS ring oscillator

2.1 Static Inverter Delay Estimation

In CMOS inverter, the output characteristics pass through few regions: cutoff ($v_{in} < V_{THn}$), saturation ($V_{THn} < v_{in} < V_{DD} - |V_{THp}|$), and linear ($v_{in} > V_{DD} - |V_{THp}|$). The voltage waveform of 5-stage inverting amplifier is plotted in Fig. 2. If we carefully noticed that these inspections are approximately between V_{THn} and $(V_{DD} - |V_{THp}|)$ [11]. The propagation delay times can be found by solving the state equation linked with output node of inverting amplifier.

$$I_{C_L} = I_{Dp,max} - I_{Dn,max} = C_L \frac{dV_{SW}}{dt} \quad (4)$$

The rising and falling slopes of node voltages can be determined from the peak current of PMOS and NMOS. Peak current can be calculated by

$$I_{Dn,max} = -\frac{C_L dV_{SW}}{dt}; \quad I_{Dp,max} = \frac{C_L dV_{SW}}{dt} \quad (5)$$

where C_L is output load capacitance of inverting amplifier, V_{SW} is output swing, $I_{Dn,max}$, and $I_{Dp,max}$ are peak currents for both MOS are in saturation region. The maximum current across NMOS and PMOS can be determined using saturated-MOS drain equation.

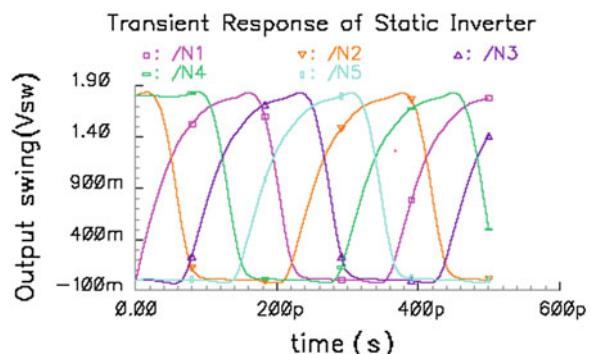
$$I_{Dn,max} = \frac{K_n}{2} \frac{W_n}{L_n} (V_{DD} - V_{THn})^2 \quad (6)$$

$$I_{Dp,max} = \frac{K_p}{2} \frac{W_p}{L_p} (V_{DD} - V_{THp})^2 \quad (7)$$

where ($W_n = W_p = W$ and $L_n = L_p = L$) and K_n and K_p are process parameters. Load capacitance can be expressed by

$$C_L = C_{ox}(W_n L_n + W_p L_p) = 2C_{ox}WL \quad (8)$$

Fig. 2 Output voltage waveform of 5-stage single-ended static inverter



The propagation delay for high-to-low (τ_{PHL}) and low-to-high (τ_{PLH}) is between V_{THn} and $V_{DD} - |V_{THp}|$, which can be expressed by

$$\tau_{PLH} = \frac{dV_{SW}C_L}{I_{Dp,max}} = \frac{4C_{ox}L^2[V_{DD} - (V_{THn} + |V_{THp}|)]}{K_p(V_{DD} - |V_{THp}|)^2} \quad (9)$$

$$\tau_{PHL} = \frac{dV_{SW}C_L}{I_{Dn,max}} = \frac{4C_{ox}L^2[V_{DD} - (V_{THn} + |V_{THp}|)]}{K_n(V_{DD} - V_{THn})^2} \quad (10)$$

To simplify above expression $|V_{THp}| = V_{THn} = V_{TH}$ (neglecting the marginally error), the propagation delay becomes as given below

$$\tau_d = \frac{\tau_{PHL} + \tau_{PLH}}{2} = \frac{2C_{ox}L^2(V_{DD} - 2V_{TH})}{(V_{DD} - V_{TH})^2} \left(\frac{1}{K_p} + \frac{1}{K_n} \right) \quad (11)$$

Now, the clock frequency can be evaluated easily from (3). From above expression, the oscillation frequency does not depend on W , but depends on L and other process parameters.

2.2 Inverter Power Consumption Optimization

The power dissipations of inverting amplifier are as follows:

- (a) **Static Power Consumption:** This power consumption is due to junction leakage current, gate tunneling leakage, and subthreshold leakage. The static power can be expressed as

$$P_{\text{static}} = V_{DD} \cdot I_{\text{Leakage}} \quad (12)$$

- (b) **Dynamic Power Consumption:** This power is due to current needed to charge and discharge the load capacitance C_L of inverting amplifier. The average power of an inverting amplifier is given as

$$P_{\text{diss}} = C_L \cdot V_{DD}^2 \cdot f \quad (13)$$

Now for N stage ring oscillator, the dynamic power dissipation can be expressed as follows:

$$P_d = N \cdot \frac{C_L V_{DD}^2}{\tau_d} \quad (14)$$

$$P_d = V_{DD}^2 \cdot \frac{W}{L} \cdot \frac{(V_{DD} - V_{TH})^2}{V_{DD} - 2V_{TH}} \cdot \frac{K_p K_n}{K_p + K_n} \quad (15)$$

In designing low-power VLSI systems, the total power dissipation is minimized for a given frequency of oscillation and supply voltage (13). The average power dissipation of a 5-stage inverting amplifier is 110 μW , and this power is the function of power supply. From above expression, static power is linearly depends on power supply but dynamic power is proportional to square of power supply. As an example, if $V_a < V_{DD}/2$, the current source dissipates more than 50 % of the total power.

3 Design Implementation

The schematic of conventional VCO is shown in Fig. 1, where the drain current controlled by bias potential (V_a). To get harmonic-like frequencies, three types of conventional designs are approached. These are (a) pseudo NMOS ring oscillator (PNRO), (b) pseudo PMOS ring oscillator (PPRO), and (c) combo ring oscillator (CRO). These three structures are discussed in the next section.

3.1 Pseudo NMOS Ring Oscillator (PNRO)

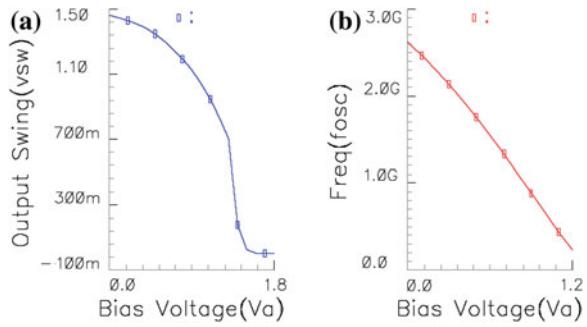
The oscillation frequency (3) strongly depends on drain current(I_{SS}) and delay stage. Pseudo NMOS inverter is a voltage-controlled inverter which is shown in Fig. 1a, where PMOS acts as a load is driven by fixed bias potential, and NMOS is driven with an input signal. In maximum cases, PMOS is in linear region so it behaves like a low-resistive path. Hence, RC time constant is low so dynamic power is low. The on resistance of PMOS written as follows:

$$R_p = \frac{L_p}{K_p W_p (V_{DD} - V_a - |V_{THp}|)} \quad (16)$$

When driver NMOS is turned ON, a constant DC current flows in the circuit. If we increase the bias voltage, PMOS enters in saturation region and frequency of oscillation given by

$$f_{osc} = \frac{K_p W_p (V_{DD} - V_a - |V_{THp}|)}{4 N C_L L_p V_{SW}} \quad (17)$$

Fig. 3 Output response for PNRO **a** output swing versus bias voltage and **b** frequency versus bias voltage



From above equation, amplitude of voltage swing across the load (V_{SW}) and oscillation frequency (f_{osc}) drops sufficiently by increasing the bias voltage (V_a). Where

$$C_L = C_{gdn} + C_{dbn} + C_{gdp} + C_{dbp} \quad (18)$$

The respective capacitance is given in [12]. The relation between frequency, output swing, and bias voltage is plotted in Fig. 3.

3.2 Pseudo PMOS Ring Oscillator (PPRO)

The basic circuit for PPRO is shown in Fig. 1b, which is just reverse of PNRO circuit. It is also a controlled inverter where NMOS acts as a load and PMOS driven with an input signal. Here, τ_{PHL} changes with bias voltage, but τ_{PLH} is constant. The on resistance of NMOS is given by

$$R_n = \frac{L_n}{K_n W_n (V_a - V_{THn})} \quad (19)$$

So, increasing bias voltage results less resistance through the discharging path. So, tail current dumped sufficiently and discharging occurs slower. Now, oscillation frequency for PPRO is given as follows:

$$f_{osc} = \frac{K_n W_n (V_a - V_{THn})}{4 N C_L L_n V_{SW}} \quad (20)$$

So, from above equation, frequency rises with rising bias voltage and output swing drops for fixed tail current. The frequency and output swing relation with bias voltage are plotted in Fig. 4.

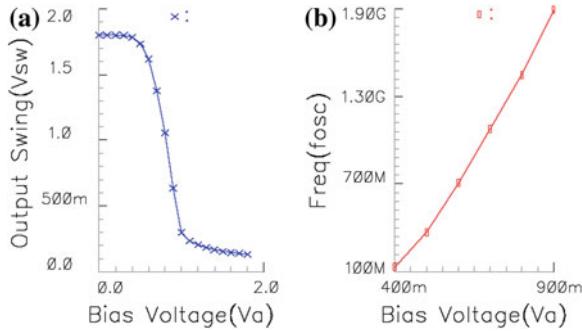


Fig. 4 Output response for PPRO **a** Output swing versus bias voltage and **b** frequency versus bias voltage

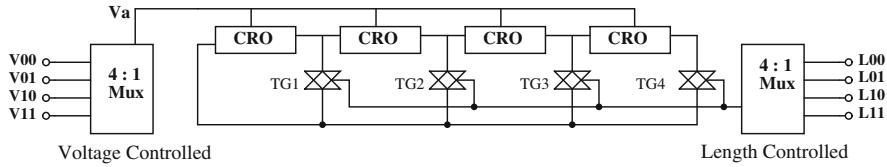


Fig. 5 Harmonic frequency generator circuit

3.3 Combo Ring Oscillator (CRO)

The CRO consists of all inverters. For CRO, we ensure that the bias voltage should lies between the minimum voltage for NMOS (V_{THn}) and maximum voltage for PMOS ($V_{DD} - |V_{THp}|$). The CRO structure is using MUX and TG which are implemented for generating harmonic frequencies depicted in Fig. 5. TG works here as an ideal switch, which controls the length of ring oscillator. Here, four TGs are implemented for designing the structure. TGs are controlled with length-controlled multiplexer. The respective lengths for respective TGs are L_{00} , L_{01} , L_{10} , and L_{11} . These four lengths signify that if TG1 is ON, then the length of the ring must be 5. If TG2, TG3, and TG4 are ON, simultaneously then the respective lengths of the ring are 7, 11, and 19, respectively (at a time single TG can be conducted).

4 Result and Discussion

The simulations carried out in cadence environment with a conventional 180 nm CMOS technology. We kept the power supply voltage (V_{DD}) at 1.8 V. To get harmonic frequencies in the sequence such as f_0 , $f_0 - \Delta f$, $f_0 - 2\Delta f$..., here a circuit is implemented using combo ring oscillator. Different bias voltages are

Table 1 Summary results of the proposed voltage-controlled ring oscillator

Analog input	Frequency from cadence (MHz)	Tested frequency (MHz)	Percentage error (%)	Average power (μ W)
$L_{00}V_{00}$	1,590	1,600	0.63	268.7
$L_{00}V_{01}$	1,500	1,500	0.00	275.5
$L_{00}V_{10}$	1,385	1,400	1.07	295.8
$L_{00}V_{11}$	1,285	1,300	1.15	342.2
$L_{01}V_{00}$	1,186	1,200	1.16	246.1
$L_{01}V_{01}$	1,105	1,100	0.45	258.2
$L_{01}V_{10}$	1,010	1,000	1.00	303.0
$L_{01}V_{11}$	940	900	4.00	360.7
$L_{10}V_{11}$	750	800	6.00	343.7
$L_{10}V_{10}$	705	700	0.70	310.1
$L_{10}V_{01}$	640	600	6.67	272.3
$L_{10}V_{00}$	520	500	4.00	219.4
$L_{11}V_{11}$	360	400	10.00	303.8
$L_{11}V_{10}$	300	300	0.00	270.5
$L_{11}V_{01}$	205	200	2.50	240.8
$L_{11}V_{00}$	108	100	8.00	232.8

passing through a 4:1 multiplexer, where control voltages are V_{00} (470 mV), V_{01} (530 mV), V_{10} (600 mV), and V_{11} (660 mV). For a particular bias voltage and for a particular TGs, 4×4 frequencies are generated. Length and voltage-controlled multiplexer decided the frequency pattern. Here, to generate harmonic frequencies, combo structure is followed for generating low as well as high frequencies. So, combo structure is generating wide range of frequency. The harmonic frequencies for different control voltages and lengths are depicted in Table 1, where average power calculation is also estimated. Here, 16 combinations of frequencies are being shown digitally.

Table 1 shows initial frequency (f_0) as 1,600 MHz and frequency deviation(Δf) is 100 MHz, also shows some error occurs due to leakage current and some passive parameters.

5 Conclusion

We have demonstrated the voltage-controlled ring oscillator structure using different inverter topology. The tuning range is 1,600–100 MHz, with supply voltage of 1.8 V and maximum power dissipation is 360.7 μ W. After further improvement, it could be less percentage of errors. This frequency range can be used in wireless communication systems.

Acknowledgments The authors would like to acknowledge the support of VLSI Laboratory at NIT Durgapur, originated from SMDP project funded by DeitY, Govt. of India.

References

1. Chen, Z.-Z., Lee, T.-C.: The design and analysis of dual-delay-path ring oscillators. *IEEE Trans. Circuits Syst. I. Reg. Papers*, **58**(3), 470–478, (2011)
2. Mandal, M.K., Sarkar, B.C.: Ring oscillator characteristics and applications. *Indian J. Pure Appl. Phys.* **48**, 136–145 (2010)
3. Min, S., Copani, T., Kiaei, S., Bakkaloglu, B.: A 90-nm CMOS 5-GHz ring-oscillator PLL with delay-discriminator-based active phase-noise cancellation. *IEEE J. Solid-State Circuits* **48** (5), 1151–1160 (2013)
4. Gupta, Nisha: Voltage controlled ring oscillator for low power noise application. *Int. J. Comput. Appl.* **14**(5), 23–27 (2011)
5. Abidi, Asad A.: Phase noise and jitter in CMOS ring oscillators. *IEEE J. Solid-State Circuits* **41**(8), 1803–1816 (2006)
6. Ramiah, H., Keat, C.W., Kanesan, K.: Design of low-phase noise, low-power ring oscillator for OC-48 application. *IETE J. Res.* **58**(5), 425–428 (2012)
7. Kim, J., Cho, S.: A time-based analog-to-digital converter using a multi-phase voltage-controlled oscillator. In: *IEEE International Symposium on Circuits and Systems*, pp. 3934–3937, May 2006
8. Aranda, M.L., Daz, O.G., Ivarez, C.R.B.: A performance comparison of CMOS voltage-controlled ring oscillators for its application to generation and distribution clock networks. *Sci. J. Circuits Syst. Sig. Process.* **2**(2), 56–66 (2013)
9. Razavi, B.: A study of phase noise in CMOS oscillators. *IEEE J. Solid-State Circuits*, **31**, 331–343 (1996)
10. Nguyen, M.H., Pham, C.K.: A wide frequency range and adjustable duty cycle CMOS ring voltage controlled oscillator. In: *IEEE International Conference on Communications and Electronics*, pp. 107–109, Aug 2010
11. Todani, R., Kumar Mal, A., Baran Maji, K.: Low frequency ring oscillator and its use in non overlapping clock generation. *IACSIT Int. J. Eng. Technol.* **4**(6), 0975–4024 (2012)
12. Docking, S., Sachdev, M.: An analytical equation for the oscillation frequency of high-frequency ring oscillator. *IEEE J. Solid-State Circuits*, **39**(3), 533–537 (2004)

Amplifier Design Optimization in CMOS

Sumalya Ghosh, Ashis Kumar Mal and Surajit Mal

Abstract A simple and realistic method is introduced to design an analog amplifier, having some design criteria, using Alpha-Power MOS law model, which becomes notable in short-channel MOSFETs (SCMs). Estimation of α , V_{th} , and k in Alpha-Power MOS law from simulation is the fundamental job to design an analog circuit using SCM, so that the simulated drain current should fit the Alpha-Power-based drain current equation. Work is done by simulation in UMC 180-nm technology. Design starts by extracting I-V value from the characteristics curve of a device NMOS using simulator. This paper also includes the variation of α , k with respect to gate voltage to minimize the design errors. Device dimension setup using the estimated value to meet the design criteria is described. Design procedures and analysis of simulated data using proposed method are briefly described and verified by designing an amplifier with resistive load. Proposed method is much more efficient, fully technology independent and free from complex mathematical expressions associated with the short-channel devices. Proposed method shows design performance quite closer and acceptable also very much suitable for initial design based on hand calculation.

Keywords Alpha-power MOS law · Analog amplifier · Device transconductance

S. Ghosh (✉) · A.K. Mal · S. Mal
National Institute of Technology, Durgapur 713209, India
e-mail: sumalyaghosh.2007@gmail.com

A.K. Mal
e-mail: toakmal@gmail.com

S. Mal
e-mail: surajit3855@gmail.com

1 Introduction

Initial design of analog circuit based on simulator and pen–paper is not an easy task due to unavailable of proper I–V relationship corresponds to designing device under sub-nanometer technology. The most simplest method that describes circuit design and behavior modeling is Shockley’s square law MOSFET model due to its simple mathematical I–V relationship. Recent, short-channel MOSFET (SCM) does not follow Shockley’s square law MOSFET model. To fit I–V curve of a SCM with an equation, as well as to analyze, the circuit design using SCM’s Sakurai proposed one model, namely Alpha-Power Law MOS Model. Mathematical expression of Sakurai’s model represents both quadratic and linear dependence of drain current (I_d) over overdrive voltage ($V_{gs} - V_{th}$). Initially, Sakurai’s model was purely empirical and mathematical, without physical background. The physical interpretation of alpha-power dependency investigated by researchers [1, 2] that shows the origin of alpha-power dependency is basically mobility degradation at high electric field. High degree of accuracy can be obtained by this model other than existing complex models such as SPICE Level 3 model [3, 4] and BSIM [5]. Once we get the I–V equation of a device properly, then the pen–paper-based design would be easier. Review [6] shows the value of alpha and has been fixed almost near about at 1.3 for SCM. Drain current error ($\Delta I_d = (I_{d,cal} - I_{d,sim})$) becomes significant in lower and upper region of V_{gs} of a SCM. So, simulator gives an erroneous result for a design using SCM. Simulated plot shows that α decreases with increase in V_{gs} not fixed throughout the V_{gs} region. Therefore, design error becomes maximum when bias voltage (V_{bias}) falls in the lower and upper range of V_{gs} . This paper proposed one method to fit the simulated data with the alpha-power law model by extracting the unknown parameter value in Alpha-Power Law, so that designers can design a circuit easily using pen–paper and circuit simulator for their initial design of interest. Based on this proposed method, the circuit simulator gives more accurate result.

First, the computation procedure of the parameter is introduced in Sect. 2. Then in Sect. 3, design procedure of an amplifier is briefly described. An example design and comparison from the existing α value, reported by Sakurai, T [6] based design has drawn in Sect. 4. Section 5 is dedicated to conclusions.

1.1 Alpha-Power Law MOS Model

The full description of Alpha-Power Law MOS model [7] in equation form is described below:

$$I_d = \begin{cases} 0 & (V_{gs} \leq V_{th} : \text{Cutoff Region}) \\ \frac{I'_{d0}}{V'_{d0}} & (V_{ds} < V'_{d0} : \text{Triode Region}) \\ I_{d0} & (V_{ds} \geq V'_{d0} : \text{Pentode Region}) \end{cases} \quad (1)$$

where

$$V'_{d0} = V_{d0} \left(\frac{V_{gs} - V_{th}}{V_{DD} - V_{th}} \right)^{\alpha/2} (= P_V (V_{gs} - V_{th})^{\alpha}) \quad (2)$$

$$I'_{d0} = I_{d0} \left(\frac{V_{gs} - V_{th}}{V_{DD} - V_{th}} \right)^{\alpha} \left(= \frac{W}{L_{eff}} P_C (V_{gs} - V_{th})^{\alpha} \right) \quad (3)$$

where V_{DD} signifies a supply voltage and P_V and P_C are parameters fully described in [7]. Here, we have assumed $\left(\frac{W}{L_{eff}} P_C \right) = k$, which is not fixed throughout the V_{gs} range. Therefore, the drain current equation becomes:

$$I_d = k (V_{gs} - V_{th})^{\alpha} \quad (4)$$

where α and k are velocity saturation index, and device transconductance, respectively.

2 Computation of Parameters

This section describes the computation procedure of unknown parameter's value in Alpha-Power Law MOS Model-based drain current equation for a particular device of width W and length L in sub-micron region. By simulating a device first, plot $I_d - V_{gs}$ characteristics of the device which is to be used for extraction of unknown device parameter corresponds to the simulated device. We have to estimate the basic unknown parameters V_{th} , k , and α related to Alpha-Power MOS Law that will clearly fit the extracted device data. Mathematically, we can say if the sum of errors between the calculated data ($I_{d,cal}$) for a particular combination of this three parameter in Alpha-Power Law and simulated data ($I_{d,sim}$) becomes zero or minimum over a certain range of V_{gs} , then the equation for this specific parameters will fit the simulated data well over the specific V_{gs} range. Therefore, this combination will give the current equation more accurately and thus describe the device I-V characteristics properly. One programming language is required to check the best combination of the specified parameters. The mathematical expression of the above procedure to extract the parameters value is given below:

$$\sum \Delta I_d = \left\{ \sum_{V_{gsi}, \alpha_i, V_{thi}, k_i}^{V_{gsf}, \alpha_f, V_{thf}, k_f} |I_{d,sim} - I_{d,cal}| \right\} \simeq 0. \quad (5)$$

In the above equation, suffixes i and f represent initial and final values, respectively, of every device unknown parameter. $I_{d,cal}$ can be calculated using the Eq. (5).

2.1 Computation Results

We have seen error (ΔI_d) becomes maximum in the range of 0–650 mV of gate voltage and with the increase in V_{gs} , error becomes smaller but average error increases with increase in V_{gs} range. Figure 1 shows the input voltage versus simulated and calculated drain current (red line for simulated and black line for calculated), simulation taken against 240/180-nm devices and calculation done for $\alpha = 1.54$, $k = 123$, $V_{th} = 0.39$ V as chosen input voltage range was 0–900 mV.

Basically, V_{bias} falls within the range of 0–900 mV; therefore, for biasing purpose, we operate an amplifier in the range of 0–900 mV of V_{gs} and have taken this range of input voltage for our design analysis. Tables 1, 2, and 3 show variation of α , k , V_{th} with V_{gs} range for various device sizes. When V_{gs} range decreases, the value of α increases also the value of k , whereas the V_{th} decreases. Simulator gives much error if we do not set the device width(W) according to α , k value corresponds to V_{gs} range, as device parameter changes with V_{gs} range not fixed through out entire region of gate voltage. Generally, V_{bias} falls in between 500 and 900 mV, so is the main reason to find the extraction of parameters in the specified range.

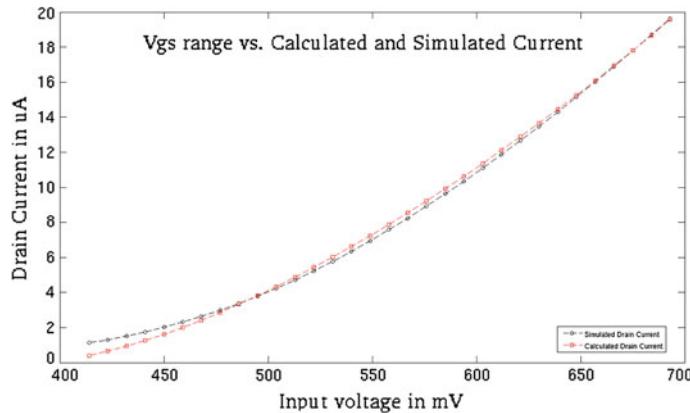


Fig. 1 Gate voltage (mV) versus drain current (μ A)

Table 1 Gate voltage (V_{gs}) versus device parameter for 240/180-nm device

Gate voltage range (V_{gs}) (mV)	α	V_{th} (V)	k (μ A/V x)	Avg. error in I_d (%)
0–900	1.54	0.39	123	7.50
0–800	1.77	0.36	136	4.60
0–700	1.98	0.34	155	3.66
0–600	2.65	0.28	225	1.31
0–500	3.05	0.26	276	0.54

Table 2 Gate voltage (V_{gs}) versus device parameter for 480/360-nm device

Gate voltage range (V_{gs}) (mV)	α	V_{th} (V)	k ($\mu\text{A}/\text{V}^x$)	Avg. error in I_d (%)
0–900	1.85	0.44	140	17.95
0–800	2.12	0.41	158	15.83
0–700	2.70	0.36	216	9.21
0–600	3.80	0.29	397	3.89
0–500	5.41	0.20	694	1.29

Table 3 Gate voltage (V_{gs}) versus device parameter for 500/500-nm device

Gate voltage range (V_{gs}) (mV)	α	V_{th} (V)	k ($\mu\text{A}/\text{V}^x$)	Avg. error in I_d (%)
0–900	1.92	0.42	108	15.35
0–800	2.17	0.39	118	11.07
0–700	2.53	0.36	145	8.01
0–600	3.54	0.29	249	2.67
0–500	5.79	0.17	596	0.79

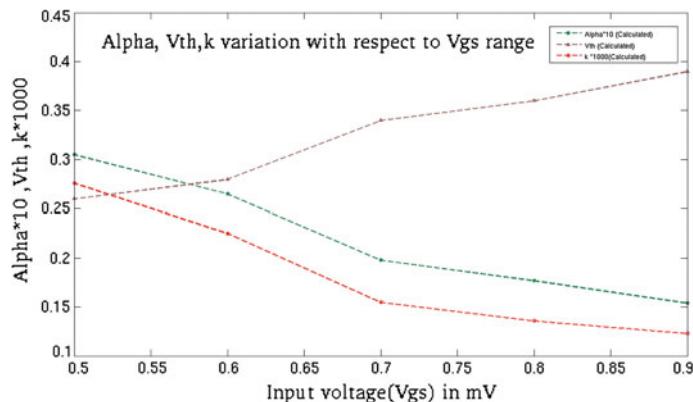
**Fig. 2** Gate voltage (V_{gs}) versus α , k , and V_{th}

Figure 2 show the actual variation of α , k , and V_{th} with respect to various range of input voltage. Green, gray, and red line shows the variation of α , V_{th} and k , respectively.

To get more accurate result, we have to estimate the device parameter for a small region of interest of gate voltage. Tables 4, 5, and 6 shows average percentage error in drain current corresponds to a particular region of interest in gate voltage. So, the simulator gives more accurate result if we calculate device width by using α and k corresponds to a region of interest in V_{gs} .

Table 4 Gate voltage (V_{gs}) versus device parameter for 240/180-nm device

Gate voltage range (V_{gs}) (mV)	α	V_{th} (V)	k ($\mu\text{A}/\text{V}^z$)	Avg. error in I_d (%)
800–900	1.21	0.47	120	0.027
700–800	1.33	0.44	122	0.052
600–700	1.62	0.39	136	0.074
500–600	2.05	0.34	172	0.085
400–500	3.8	0.2	392	0.307

Table 5 Gate voltage (V_{gs}) versus device parameter for 480/360-nm device

Gate voltage range (V_{gs}) (mV)	α	V_{th} (V)	k ($\mu\text{A}/\text{V}^z$)	Avg. error in I_d (%)
800–900	1.62	0.48	135	0.068
700–800	1.87	0.44	144	0.085
600–700	2	0.43	159	0.072
500–600	4.13	0.27	455	0.64
400–500	5.62	0.19	743	1.13

Table 6 Gate voltage (V_{gs}) versus device parameter for 500/500-nm device

Gate voltage range (V_{gs}) (mV)	α	V_{th} (V)	k ($\mu\text{A}/\text{V}^z$)	Avg. error in I_d (%)
800–900	1.74	0.45	105	0.0303
700–800	1.64	0.47	104	0.0784
600–700	1.97	0.42	115	0.0712
500–600	2.75	0.35	178	0.1830
400–500	6.27	0.15	705	0.3376

3 Design Procedure of an Amplifier

This section describes the design procedure steps in details using our proposed method. The design steps are enlisted below:

1. Estimation of unknown device parameter in Alpha-Power Law MOS model of a designing device of width W and length L is the first step to design an amplifier. By implementing the mathematical concept described in Eq. (4) with C-programming or matlab coding, designers can estimate the device parameter value.
2. Calculate the load resistance (R_{load}) to satisfy the bias current (I_{bias}) using the following equation.

$$R_{load} = \frac{V_{DD} - V_{DS}}{I_{bias}} \quad (6)$$

3. Again, the gain of an amplifier with passive load (R_{load}) can be described as below:

$$g_m R_{\text{load}} = A_v \quad (7)$$

Equation (7) is used to calculate the g_m of the device corresponds to gain (A_v) and bias current (I_{bias}), where g_m and A_v are the transconductance and gain, respectively. Calculated g_m is required to set the device dimension to meet expected amplifier's design criteria.

4. Device dimension has to be increase, to meet design criteria can be determined using the equation mentioned below:

$$r = \left(\frac{g_m}{\alpha}\right)^{\alpha} I_d^{(1-\alpha)} \left(\frac{1}{k}\right) \quad (8)$$

where, r is a device dimension scaling factor.

5. Required width (W') and length (L') of the device will be:

$$\left(\frac{W'}{L'}\right) = r \left(\frac{W}{L}\right) \quad (9)$$

where W and L are the width and length of the simulated device, W' and L' are the required width and length of the designing device.

6. Check the biasing voltage (V_{bias}) corresponds to bias current (I_{bias}) and falling region of V_{gs} where V_{bias} falls. Therefore modified device dimension, according to α , k corresponds to the falling region of V_{bias} in V_{gs} , can be calculated using Eqs. (8) and (9).
7. After modifying the device dimension repeat steps 4, 5 and 6, until V_{bias} tightly fit the falling region of V_{gs} .
8. To fulfill the required criteria, drain to source resistance (r_o) has to be included as it is not zero for a bias voltage V_{bias} . After including r_o , simulation gives accurate result where design error within the range of 10 %.
9. Actual required g_m includes r_o to fulfill the required gain and other criteria. So, $(g_m)_{\text{req.}}$ will be:

$$(g_m)_{\text{req.}} = \left(\frac{A_v}{R_{\text{load}}} + \frac{A_v}{r_o} \right) \quad (10)$$

10. So, the final step is to set the device dimension $\left(\frac{W'}{L'}\right)$ correspond to $(g_m)_{\text{req.}}$ using Eqs. (8) and (9). After that repeat steps 6 and 7 to meet the required design specification.

4 Example Design

This section describes an example design in details of an amplifier of specification say, bias current(I_{bias}) = 20 μ A, gain(A_v) = 10. Say, V_{DS} = 900 mV. Figure 3 shows a simple CS amplifier with resistive load. So the required load resistance becomes (R_{load}) = 45 k Ω . Using the procedure described in Sect. 3, the calculated width and gain becomes 2.35 and 9.1 μ m, respectively. The required g_m was 312 μ A/V; after simulation, g_m value becomes 303 μ A/V. Table 7, shows the simulated data and actual data.

Initial device width was 240, 480, and 500 nm for length 180, 360, and 500 nm, respectively. For higher gain, device length has to be increased more than 1 μ m; otherwise, the transconductance does not meet the required $g_{m,req}$. Table 8 shows comparison of amplifier gain (A_v) (actual and simulated), required g_m ($g_{m,req}$) and simulated g_m ($g_{m,sim}$), designed using our proposed method-based and reported by A JSSC Classic paper [6]-based α -value. Length of the device was 180 nm for gain 10 and 360 nm for 15 in comparison table. So, from Table 8, we can conclude that proposed method-based design error lies within 10 % whereas that lies within 30 % using data reported by A JSSC Classic paper [6].

Fig. 3 CS amplifier with resistive load

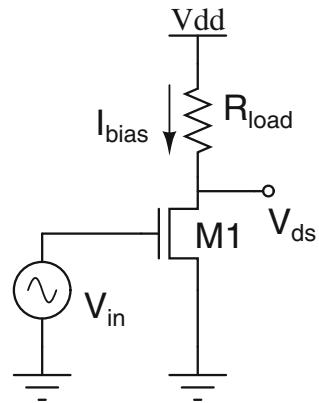


Table 7 Device dimension versus device parameter

Device length (nm)	Gain (actual)	Gain (simulated)	$g_{m,req.}$ (μ A/V)	$g_{m,sim}$ (μ A/V)
360	15	14.78	403	412
500	15	14.78	403	412
180	10	9.1	313	304
500	10	10	268	278
180	5	4.78	142	138
360	5	5	125	127

Table 8 Comparison of gain (A_v) and g_m of an amplifier based on reported and proposed method-based α -value

α -value from	$A_{v,actual}$	$A_{v,sim}$	$g_{m,req.}$ ($\mu\text{A/V}$)	$g_{m,sim}$ ($\mu\text{A/V}$)
Proposed method	10	9.1	313	304
Ref. [6]	10	7.56	307	240
Proposed method	15	14.78	403	412
Ref. [6]	15	10.13	409	280

4.1 CS Amplifier Stage with Current Source Load

Design of CS amplifier with PMOS current source load using Alpha-Power Law is presented in this section. Consider the CS amplifier with current source load shown in Fig. 4, where MOSFET M1 is the common-source component of the amplifier, while MOSFET M2 is the current source load. Required bias voltage has to be provided from the biasing circuit. The load resistance variation with respect to bias current (I_{bias}) is the main concern in our design consideration. Here, R_{load} becomes $\left(\frac{1}{g_{ds2}}\right)$ for a particular bias current. Figure 5 shows the variation of load resistance of a MOSFET with respect to bias current (I_{bias}). From this curve, we can easily obtain the output resistance (r_o) of a device corresponds to a particular bias current (I_{bias}). The current source load-based CS amplifier design procedure is somehow same as CS amplifier with passive load. As active load offers largest possible load resistance and thus largest gain, therefore, the required g_m becomes lower than the passive load-based design. If the resistance looking into the source of M2 is $\left(\frac{1}{g_{ds2}}\right)$, and the resistance of M1 is $\left(\frac{1}{g_{ds1}}\right)$, then the effective load resistance acting on the drain of M1

Fig. 4 Load resistance versus device length

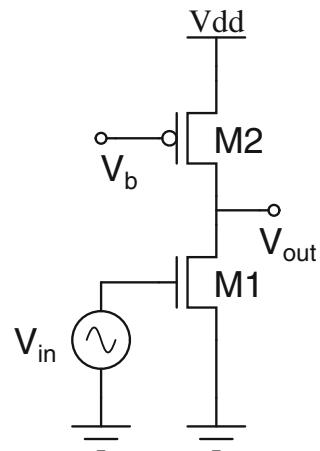
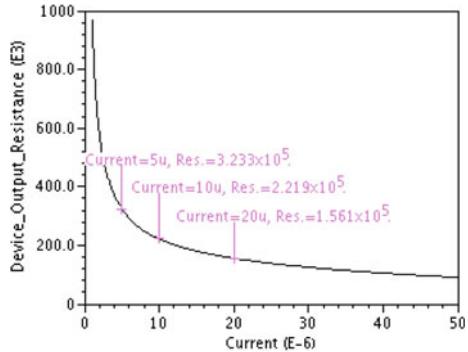


Fig. 5 Variation of device output resistance with respect to bias current of 240/180-nm device

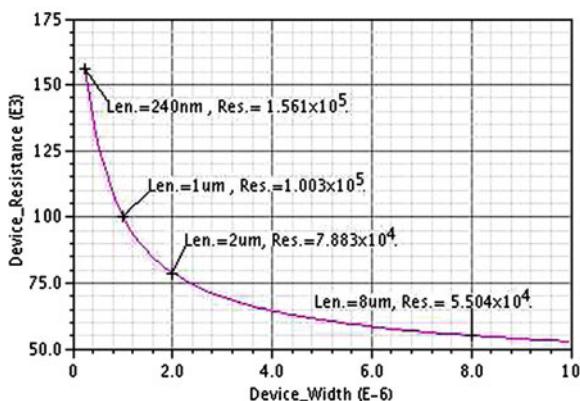


becomes $r_{o,eq} = \left(\frac{1}{g_{ds2}} \parallel \frac{1}{g_{ds1}} \right)$. For this case, the required $g_{m,req}$ becomes $\approx \left(\frac{A_v}{r_{o,eq}} \right)$. Table 9 shows the measured and calculated data for various gain and corresponding device length using our proposed method. Error lies within 20 % range. The channel length modulation factor (λ) reduces the design error of an amplifier. Therefore, we have to include (λ) to reduce the design error, though it is not a dominant factor. The deviation from the expected value is due to the fact that we neglected the effect r_o with device width (W) in gain calculation ($A_v \propto R_L \parallel r_o$). Figure 6 shows the r_o versus

Table 9 Device dimension versus device parameter

Device length (nm)	Gain (actual)	Gain (simulated)	$g_{m,req}$ ($\mu\text{A/V}$)	$g_{m,sim}$ ($\mu\text{A/V}$)
180	20	15.70	222	218
360	20	18.35	154	153
500	20	18.96	136	136
180	25	19.29	358	333
360	25	21.46	192	192
500	25	24.05	190	190

Fig. 6 Device width (W) versus device output resistance (r_o)



device width (W) of a device. We have seen if the device width (calculated, until such trade-offs between V_{th} , α , and k with respect to gate voltage region do not appear) does not differ much than the previously calculated value, then the effective difference of r_o on gain calculation is negligible; otherwise, we have to consider the variation of r_o with device width in our design.

5 Conclusion

A new design method using Alpha-Power MOS Law model is introduced to overcome the design problem issue such as device width and g_m . This new approach can express the salient features of the SCM I-V characteristics. The method is very much simple and suitable for circuit design and behavior analysis. Useful expressions for required device width corresponds to design criteria are derived. It has been shown that proposed method is fully technology independent and free from complex mathematical expressions associated with the devices. This simple method shows that the design performance is quite closer and is very much suitable for initial design based on hand calculation. In this work, an amplifier with resistive also with PMOS current source load is designed for a given specification is described to validate the method. All the designs have been simulated in UMC 180-nm CMOS technology process.

References

1. Bowman, K.A., Austin, B.L., Eble, J.C., Tang, X., Meindl, J.D.: A physical alpha-power law MOSFET model. *IEEE J. Solid-State Circuits* **34**(10), 1410–1414 (1999)
2. Im, H., Song, M., Hiramoto, T., Sakurai, T.: Physical insight into fractional power dependence of saturation current on gate voltage in advanced short-channel MOSFETs (alpha-power law model). In: Proceedings of ISLPED 2002, pp. 13–18. Aug 2002
3. Vladimirescu, A., Liu, S.: UC Berkeley ERL Memo M80/7. Oct 1980
4. Quarles, T., Newton, A.R., Pederson, D.O., Sangiovanni-Vincentelli, A.: SPICE 3B1 User's Guide. EECS University of California, Berkeley. (1988)
5. Sheu, B.J.: UCB/ERL Memo M85/85. Oct 1985
6. A JSSC Classic paper: The simple Model of CMOS Drain Current, vol. 9, No. 4, Oct 2004
7. Sakurai, T., Newton, A.R.: Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas. *IEEE J. Solid-State Circuits* **25**(2), 584–594 (1990)

An Application of ANFIS-Based Intelligence Technique for Predicting Tool Wear in Milling

Shibendu Shekhar Roy

Abstract In this work, an attempt has been made to design an intelligence technique-based expert system using adaptive neuro-fuzzy inference system (ANFIS) for predicting tool wear in milling operation. An artificial neural network is used for designing an optimized fuzzy logic system, so that the tool wear can be modeled for a set of input cutting parameters, namely feed rate, depth of cut, and cutting force. The proposed method uses two different learning approaches, namely back-propagation gradient descent method alone and hybrid method (i.e., combination of the least squares method and back-propagation algorithm) for training of first-order Sugeno-type fuzzy system. The predicted tool wear values derived from proposed ANFIS were compared with the experimental data.

Keywords ANFIS · Tool wear · Milling

1 Introduction

The concept of tool condition monitoring has gained considerable importance in the manufacturing industry. This is mainly attributed to the transformation of the manufacturing environment from manually operated production machines to CNC machine tools and the highly automated CNC machining center. For modern machine tools, 20 % of the downtime is attributed to tool failure, resulting in reduced productivity and economic losses [1]. So in machining processes, tool condition is of vital importance, as it affects the quality of the product and the efficiency of the process. Different types of sensor have been used to detect tool wear. Yao et al. [2] and Lee et al. [3] used acoustic emission in the development of a tool wear monitoring system. Mehta et al. [4] employed an accelerometer to

S.S. Roy (✉)

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

e-mail: ssroy99@yahoo.com

monitor tool wear and concluded that the frequency increases as the tool deteriorates. Konig et al. [5] used cutting force to predict tool wear and found that there is a positive correlation between cutting tools and tool deterioration. Lin and Yang [6] proposed a tool wear monitoring system in face milling operations using a force coefficient. Artificial neural network was used for predicting tool wear in milling by Chen et al. [7]. Wang et al. [8] presented a tool wear monitoring system using a fuzzy logic model called fuzzy clustering. However, this model is complicated and there is no current evidence of model validation. Moreover, the proper selection of the number, the type, and the parameters of the fuzzy membership functions and rules are crucial for achieving the desired performance. Yet, it has been often done through trial and error. This fact highlights the significance of fuzzy logic system tuning. To achieve this, adaptive neuro-fuzzy approach is used to model tool wear in milling operation as milling is one of the most common metal removal operations used in manufacturing industry. In this work, artificial neural network is used for designing an optimized fuzzy logic system, so that the tool wear can be modeled for a set of input parameters, namely feed rate, depth of cut, and cutting force.

2 Proposed Adaptive Neuro-Fuzzy Inference System

2.1 Fuzzy Logic System

Fuzzy set theory was developed by Prof. L.A. Zadeh [9] to capture the imprecise modes of reasoning employed in an environment characterized by uncertainty and vagueness. A fuzzy inference system employing fuzzy ‘if-then’ rules can model the qualitative aspects of human knowledge and reasoning processes without employing precise quantitative analysis. This fuzzy modeling first explored systematically by Takagi and Sugeno [10]. A fuzzy inference system is composed of a rule base, containing a number of fuzzy if-then rules; a database, which defines the membership functions of fuzzy sets used in fuzzy rules; a decision-making unit which performs inference operations on the rules; and a defuzzification interface, which transforms fuzzy results into a crisp output. The integration of neural network into fuzzy logic system makes it possible to learn from data prior.

2.2 Adaptive Neuro-Fuzzy System

The adaptive neural-fuzzy system is a Sugeno fuzzy model put in the framework of adaptive systems to facilitate learning and adaptation [10]. Such framework makes the adaptive neuro-fuzzy modeling more systematic and less reliant on expert

knowledge. To present the adaptive neural-fuzzy system architecture, let us consider two fuzzy if-then rules based on a first-order Sugeno model:

Rule 1: If (v is V_1) and (d is D_1), then $f_1 = p_1v + q_1d + t_1$

Rule 2: If (v is V_2) and (d is D_2), then $f_2 = p_2v + q_2d + t_2$

where V_i and D_i are the fuzzy sets, f_i is the outputs within the fuzzy region specified by the fuzzy rule, and p_i , q_i , and t_i are the design parameters that are determined during the training process. The entire system architecture consists of five layers, namely the fuzzy layer, product layer, normalized layer, de-fuzzy layer, and total output layer.

Fuzzy layer: Every node in this layer is an adjustable node, with node function as

$$O_{1,i} = \mu_{V_i}(v) \quad i = 1, 2 \quad (1)$$

$$O_{1,i} = \mu_{D_i}(d) \quad i = 3, 4 \quad (2)$$

where $\mu_{V_i}(v)$ and $\mu_{D_i}(d)$ can adopt any fuzzy membership function. For example, if the Gaussian membership function is employed, $\mu_{V_i}(v)$ is given by:

$$\mu_{V_i}(v) = e^{\frac{-(v-c_i)^2}{2\sigma_i^2}} \quad (3)$$

where $\{\sigma_i, c_i\}$ represents the parameter set. It is significant that if the values of these parameters set change, the Gaussian function will be changed accordingly. The parameters in this layer are named as ‘premise parameters.’

Product layer: Every node in this layer is a fixed node with node function to be multiplied by input signals to serve as output signal

$$O_{2,i} = w_i = \mu_{V_i}(v) \times \mu_{D_i}(d) \quad i = 1, 2 \quad (4)$$

The output signal w_i means the firing strengths of a rule.

Normalized layer: Every node in this layer is a fixed node with node function to normalize firing strength by calculating the ratio of this node firing strength to the sum of the firing strength:

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad i = 1, 2 \quad (5)$$

De-fuzzy layer: Every node in this layer is an adjustable node with node function as

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i(p_i v + q_i d + t_i) \quad i = 1, 2 \quad (6)$$

where p_i , q_i , and t_i are parameter set which is referred as the ‘consequent parameters.’

Total output layer: Every node in this layer is a fixed node with node function to compute the overall output by

$$O_{5,i} = \sum_{i=1}^2 \bar{w}_i f_i \quad (7)$$

2.3 Learning Algorithm of Adaptive Neuro-Fuzzy System

It can be observed that there are two adaptive layers in this architecture, namely the first layer and the fourth layer. In the first layer, there are two modifiable parameters $\{\sigma_i, c_i\}$, which are related to the input membership functions. In the fourth layer, there are also three modifiable parameters $\{p_i, q_i, t_i\}$, pertaining to the first-order polynomial. The task of the learning algorithm for this architecture is to tune all the modifiable parameters, namely $\{\sigma_i, c_i\}$ and $\{p_i, q_i, t_i\}$, to make the adaptive neural-fuzzy system output match the training data. The learning algorithm may consist of either back-propagation learning algorithm or hybrid learning algorithm.

Approach 1:

In this approach, the basic learning method is back-propagation algorithm which is an error-based supervised learning algorithm. It employs an external reference signal, which acts like a teacher and generates an error signal by comparing the reference with the obtained response. Based on error signal, the network modifies the design parameters to improve the system performance. It uses gradient descent method to update the parameters.

Approach 2:

The hybrid learning algorithm applies a combination of least square estimator (LSE) and gradient descent method for training fuzzy logic system membership function parameters to emulate a given training data set. The LSE method is used to optimize the consequent parameters with the premise parameters fixed. Once the optimal consequent parameters are found, the gradient descent method is used to adjust optimally the premise parameters corresponding to the fuzzy sets in input domain. The output of the adaptive neuro-fuzzy system is calculated by employing the consequent parameters. The output error is used to adapt the premise parameters by means of a standard gradient descent method.

3 Results and Discussion

A training database with regard to machining parameters and tool wear is essential to train the fuzzy system for modeling and predicting tool wear. A number of milling experiments were carried out [7] on Fadal machine using 3-insert with a diameter 1.25 in. mill and 1,018 steel workpiece. The worn tools, which were

ground artificially, were used in the experimental runs. The experimental design for collecting data was set up as the following combination of machining cuts: feed rate at 5.0, 7.0, 9.0, 11.0, and 13.0 in./min; depth of cut at 0.02, 0.03, 0.04, 0.05, and 0.06 in.; flank wear at 0.25, 0.35, 0.45, and 0.55 mm. Based on above combinations, 100 milling experiments were performed. The average peak cutting force in the y-direction collected through the experimental runs and unit for cutting force is measured in V. These sets of 100 data, as listed in Table 1, were used for training of the fuzzy system by both Approach 1 and Approach 2 as mentioned in Sect. 2.3. This training adjusts the membership function parameters. In general, this type of model works well if the training data presented are fully representative of the features of the data that the trained fuzzy system is intended to model. This is not always the case; however, in some cases, data are collected using noisy measurements, and the training data cannot be representative of all the features of the data

Table 1 Experimental data [7] for training

Feed rate (in./min)	Depth of cut (in.)	Tool wear (mm)			
		0.25	0.35	0.45	0.55
		Cutting force (V)			
5.0	0.02	0.31	0.23	0.11	0.08
	0.03	0.39	0.29	0.18	0.11
	0.04	0.45	0.35	0.21	0.17
	0.05	0.56	0.41	0.29	0.25
	0.06	0.59	0.45	0.37	0.29
7.0	0.02	0.39	0.28	0.16	0.13
	0.03	0.48	0.37	0.27	0.20
	0.04	0.56	0.42	0.35	0.29
	0.05	0.64	0.47	0.44	0.38
	0.06	0.67	0.54	0.49	0.43
9.0	0.02	0.45	0.31	0.21	0.15
	0.03	0.52	0.40	0.33	0.25
	0.04	0.62	0.43	0.44	0.35
	0.05	0.71	0.56	0.52	0.43
	0.06	0.76	0.62	0.59	0.54
11.0	0.02	0.48	0.40	0.25	0.19
	0.03	0.62	0.51	0.36	0.28
	0.04	0.76	0.64	0.50	0.39
	0.05	0.81	0.71	0.62	0.47
	0.06	0.87	0.76	0.72	0.62
13.0	0.02	0.61	0.50	0.31	0.26
	0.03	0.71	0.60	0.44	0.35
	0.04	0.78	0.72	0.58	0.47
	0.05	0.95	0.78	0.71	0.60
	0.06	1.00	0.91	0.86	0.72

that will be presented to the model. So, model validation is needed to cross validate the fuzzy system using testing data set. The testing data set is useful in checking the generalization capability of the resulting fuzzy system. That is why another 9 sets [7] were then used for testing after the training was completed to verify the accuracy of the predicted values of tool wear.

The fuzzy expressions for feed rate are LF: low feed, MF: medium feed, HF: high feed; for depth of cut are SA: shallow, MD: medium deep, DE: deep; and for cutting force are LFR: low force; MFR: medium force; HFR: high force. There are a total of 27 fuzzy rules in the architecture. Thus, a typical i th rule of the fuzzy system will be as follows:

If feed rate (f_i) is high feed (HF), depth of cut (d_i) is deep (DE), and cutting force (v_i) is medium force (MFR), then tool wear is $(p_i f_i + q_i d_i + r_i v_i + c_i)$, where p_i , q_i , r_i , and c_i are the design parameters.

During training in adaptive neural-fuzzy system, 100 sets of experimental data were used to conduct 10,000 cycles and 200 cycles of learning for Approach 1 and Approach 2, respectively. Adaptive neuro-fuzzy system learning scenario for predicting tool wear is as follows:

Number of nodes = 78, Number of linear parameters = 108, Number of non-linear parameters = 18.

Table 2 compares the predicted values and experimental values [7] of tool wear after training by using Approach 1 and Approach 2 for some of the training cases. In almost all the cases, prediction of tool wear using Approach 2 (average error: 2.32 % approx.) is better than Approach 1 (average error: 8.57 % approx.).

Table 2 Tool wear predicted by Approach 1, Approach 2, and its comparison with experimental results [7] for some of the training cases

Feed rate (in./min)	Depth of cut (in.)	Cutting force (V)	Tool wear (mm) experimental result	Predicted value			
				Approach 1		Approach 2	
				Tool wear (mm)	Abs. % error	Tool wear (mm)	Abs. % error
5	0.03	0.11	0.55	0.5189	5.6545	0.5598	1.7818
5	0.06	0.45	0.35	0.3183	9.0571	0.3596	2.7428
7	0.03	0.27	0.45	0.4756	5.6889	0.4521	0.4667
7	0.05	0.64	0.25	0.2693	7.720	0.2538	1.520
7	0.06	0.67	0.25	0.2403	3.880	0.2501	0.040
9	0.02	0.15	0.55	0.5358	2.5818	0.5553	0.9636
9	0.06	0.76	0.25	0.2562	2.480	0.2485	0.600
11	0.02	0.19	0.55	0.5714	3.8909	0.5552	0.9454
11	0.04	0.5	0.45	0.4451	1.0888	0.4532	0.7111
11	0.06	0.87	0.25	0.2595	3.800	0.2488	0.480
13	0.03	0.35	0.55	0.5248	4.5818	0.5541	0.7454
13	0.05	0.71	0.45	0.4225	6.1111	0.4477	0.5111

Table 3 Tool wear predicted by Approach 1, Approach 2, and its comparison with experimental results [7] for some of the test cases

Feed rate (in./min)	Depth of cut (in.)	Cutting force (V)	Tool wear (mm) experimental result	Predicted value			
				Approach 1		Approach 2	
				Tool wear (mm)	Abs. % error	Tool wear (mm)	Abs. % error
7.0	0.02	0.15942	0.45	0.5328	18.400	0.4738	5.2889
11.0	0.03	0.280193	0.55	0.5282	3.9636	0.5447	0.9636
7.0	0.05	0.47343	0.35	0.377	7.7143	0.3707	5.9143
5.0	0.04	0.347826	0.35	0.373	6.5714	0.3438	1.7714
11.0	0.06	0.874396	0.25	0.2566	2.6400	0.2444	2.2400
9.0	0.04	0.439614	0.45	0.4288	4.7111	0.395	12.222
13.0	0.02	0.497585	0.35	0.4643	32.657	0.3454	1.3143
7.0	0.06	0.536232	0.35	0.3484	0.4571	0.3491	0.2571
11.0	0.02	0.188406	0.55	0.5722	4.0364	0.5588	1.6001

Table 3 shows the comparative study between predicted tool wear using approach 1, approach 2, and experimental results obtained by Chen [7] for some of the test cases. Simulation result shows that the average error of the modeling of tool wear is approximately 3.51 and 9.01 % when Approach 2 and Approach 1 are used, respectively, whereas average error of ANN predicted model [7] is around 10.47 %. Approach 2 has been performed better than Approach 1, in most of the cases, for making prediction of tool wear in milling operation. It may happen because back-propagation algorithm is never assured of finding the global minimum. The error surface may have many local minima, so it may get stuck during the learning process on flat or near flat regions of the error surface. Results presented in Table 3 also show that the Approach 2 outperforms the ANN-based method in terms of test accuracy of tool wear.

4 Conclusions

The conclusions that were drawn from this study can be summarized in the following points: An adaptive neuro-fuzzy inference system (ANFIS) based on first-order Takagi–Sugeno fuzzy inference system is used to predict tool wear in milling operation. Simulation results show Approach 2 (i.e., hybrid learning of fuzzy system) performs better than approach 1 (i.e., back-propagation-based learning) and ANN-based predictive model. The hybrid learning algorithm converges much faster than pure back-propagation learning algorithm. The ability of the ANFIS for predicting tool wear of a machining process before carrying out actual experiment will help us to develop an intelligent manufacturing system.

References

1. Prickett, P.W., Johns, C.: An overview of approaches to end milling tool monitoring. *Int. J. Mach. Tools Manuf.* **39**, 105–122 (1999)
2. Yao, Y., Li, X., Yuan, Z.: Tool wear detection with fuzzy classification and wavelet fuzzy neural network. *Int. J. Mach. Tools Manuf.* **39**, 1525–1538 (1999)
3. Lee, L.C., Lee, K.S., Gan, C.S.: On the correlation between dynamic cutting force and tool wear. *Int. J. Mach. Tools Manuf.* **29**, 295–303 (1989)
4. Metha, N.K., Pandey, P.C., Chakravarti, G.: An investigation of tool wear and the vibration spectrum in milling. *Wear* **91**, 219–234 (1983)
5. Konig, W., Langhammer, K., Schemmel, H.U., Th, R.W.: Correlation between cutting force components and tool wear. *Ann. CIRP* **21**, 19–25 (1972)
6. Lin, S.C., Yang, R.J.: Force-based model for tool wear monitoring in face milling. *Int. J. Mach. Tools Manuf.* **35**, 1201–1211 (1995)
7. Chen, J.C., Chen, J.C.: An artificial neural networks based in-process tool wear prediction system in milling operation. *Int. J. Adv. Manuf. Tech.* **25**, 427–434 (2005)
8. Wang, Z., Lawrenz, W., Rao, R.B.K.N., Hope, T.: Featured filtered fuzzy clustering for condition monitoring of tool wear. *J. Intell. Manuf.* **7**, 13–22 (1996)
9. Klir G.J., Yuan B.: Fuzzy sets and fuzzy logic: theory and applications. Prentice-Hall of India Private Limited, Delhi (2001)
10. Roger Jang, J.S., Sun, C.T.: Neuro-fuzzy modeling and control. *Proc. IEEE* **83**, 378–404 (1995)

A High-performance Elliptic Curve Cryptographic Processor for FPGA Platform

V.R. Venkatasubramani, N. Premkumar, K. Vignesh and S. Rajaram

Abstract In this paper, an elliptic curve crypto processor (ECCP) is proposed using an improved quad Itoh–Tsujii algorithm, as primitive, on field-programmable gate arrays (FPGAs) platform for binary fields generated by irreducible trinomials. Efficiency is obtained by eliminating the precomputation steps required in conventional quad Itoh–Tsujii algorithm scheme. Experimental results show that the proposed ECCP architecture has better area-time product compared to existing techniques .

Keywords Elliptic curve cryptography (ECC) · Itoh–Tsujii algorithm (ITA) · Field-programmable gate arrays (FPGAs) · Lookup table (LUT)

1 Introduction

The use of elliptic curves in public key cryptography was first proposed independently by Miller [1] and Koblitz [2]. The main advantage of elliptic curve cryptography (ECC) over conventional asymmetric crypto systems is the increased security offered with smaller key sizes. For example, a 256-bit key in ECC produces the same level of security as a 3,072-bit RSA key. In elliptic curve cryptosystem, main operations such as key agreement, signature generation, signing, and verification involve scalar multiplication. The speed of scalar multiplication plays an important role in the efficiency of whole system. Efficiency of scalar multiplication also significantly depends on the underlying finite field primitives such as field inversion and multiplication. Multiplicative inverse computation is the most complex operation. The common algorithms for finding the multiplicative inverse are the extended Euclidean algorithms (EEA) and the Itoh-Tsujii algorithm (ITA) [3].

V.R. Venkatasubramani (✉) · N. Premkumar · K. Vignesh · S. Rajaram
Department of ECE, Thiagarajar College of Engineering, Madurai 625-015, India
e-mail: venthiru@tce.edu

Generally, the EEA and its variants, the binary EEA, and Montgomery inverse algorithms [4] result in compact hardware implementations, while the ITA is faster.

In [5], a pipelined ECC processor is developed which uses a combined algorithm to perform point doubling and point addition. This is the fastest reported in literature. However, the seven-stage pipeline used has huge area requirements. In [6], high performance is attained by focusing on efficient implementations of the finite field primitives. The improved quad Itoh-Tsujii inversion algorithm utilizes FPGA resources efficiently. However, the overhead of the quad ITA is the need to precompute a^3 . Also, it is required to store this precomputed value and intermediate results. In this paper, an improved quad Itoh-Tsujii inverse algorithm is used that takes directly into the iteration steps of the algorithm, and therefore, it is not required to precompute a^3 . This saves multiplication, register storage, and clock cycle. The multiplier used Hybrid Karatsuba Multiplier (HKM) algorithm. It is a combinational circuit that efficiently utilizes the FPGA resources. An ECCP is built by integrating the improved QITA [7] inverse module and conventional HKM as finite field primitives.

2 Improved Quad Itoh–Tsujii Algorithm

To eliminate precomputation steps and intermediate storage register from QITA architecture, an improved quad ITA is used:

Input: $a \in GF(2m)$, $m = 233$, $(m - 1) = 232$
 Brauer addition chain $U = \{1, 2, 3, 6, 7, 14, 28, 29, 58, 116\}$
 Output: The Multiplicative inverse a^{-1}

```

begin
  l' = length(u)
  βu1 = a
  for each ui ∈ (2 ≤ i ≤ l') do
    p = ui-1
    q = ui - ui-1
    βui = βq * βp4q

```

End

return $a^{-1} = [\beta_{l'}(a)]^{2(2^n-1)}$

End

The first step in the computation of $a - 1$ is the determination of (β_{u_2}) . This normally takes two clock cycles because of the two combinational multiplications. In the improved quad ITA, precomputation steps are eliminated. Therefore, initial

two multiplications are not required. However, final step requires one multiplication and one quad operation that can be performed in one clock cycle. By eliminating precomputation of a^3 , the used design does not require extra intermediate storage register bank.

3 ECC Architecture and Its Design

The processor implements the double and add algorithm. The processor shown in Fig. 1 is capable of doing the elliptic curve operations of point addition and point doubling.

Point doubling is done at every iteration of the loop in double and add algorithm, while point addition is done for every bit set to one in the binary expansion of the scalar input k . The output produced as a result of the scalar multiplication is the product kP . Here, P is the basepoint of the curve and is stored in the ROM in its affine form. Control signals generated every clock cycle depending on the elliptic curve operation control the data flow and the computation done. During the initialization phase, the curve constant b and the basepoint P are loaded from the ROM into the registers after which there are two computational phases. The first phase multiplies the scalar k to the basepoint P . The result produced by this phase is in projective coordinates. The second phase of the computation converts the projective point result of the first phase into the affine point kP . The second phase mainly involves an inverse computation. The inverse is computed using the improved quad Itoh-Tsujii inverse algorithm.

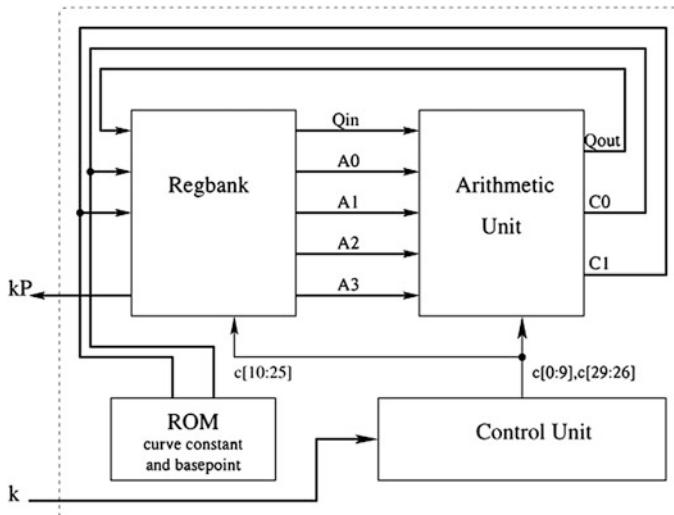


Fig. 1 Block diagram of elliptic curve crypto processor

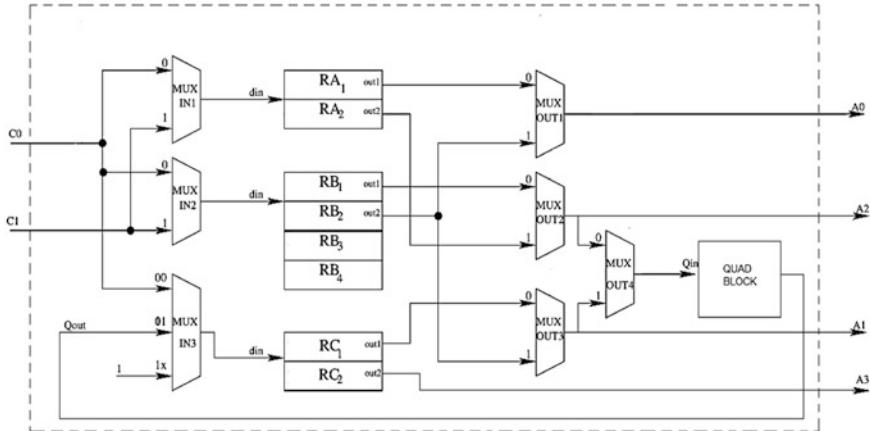


Fig. 2 Register file for ECCP in binary field

3.1 Register Module

The register file as shown in Fig. 2 consists of eight 233-bit registers. The registers are used to store the results of the computations done at every clock cycle. The registers are dual ported and arranged in three banks RA, RB, and RC. The dual-ported RAM allows asynchronous reads on the lines OUT1 and OUT2 corresponding to the address on the address lines ad1 and ad2, respectively. A synchronous write of the data on din is done to the location addressed by ad1. On the FPGA, the registers are implemented as distributed RAM.

At every clock cycle, the register file is capable of delivering five operands on buses A_0 , A_1 , A_2 , A_3 , and Q_{in} to the ALU and able to store three results from buses C_0 , C_1 , and Q_{out} . The inputs to the register file are either the ALU outputs, the curve constant b , or the basepoint $P = (P_x, P_y)$.

3.2 Architecture of ALU

The arithmetic unit shown in Fig. 3 is built using finite field arithmetic circuits and organized for efficient implementation of point addition and point doubling in LD ordinates.

The ALU has 5 inputs (A_0 – A_3 and Q_{in}) and 3 outputs (C_0 , C_1 , and Q_{out}). The main components of the ALU are a quadblock and a multiplier. The multiplier is based on the Hybrid Karatsuba algorithm. It is used in both phases (during the scalar multiplication phase and conversion to affine coordinate phase) of the computation. The ALU has several adders and squarer circuits. These circuits are small compared to the multiplier and the quadblock and therefore contribute marginally to the overall

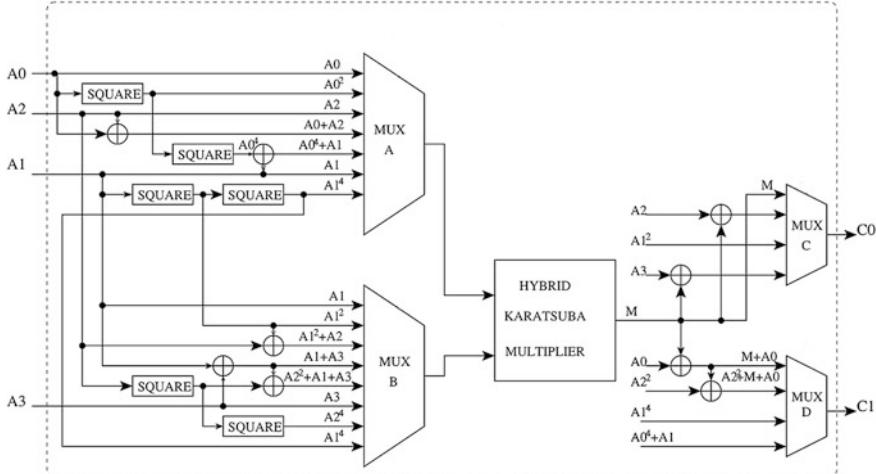


Fig. 3 Arithmetic logic unit

area and latency of the processor. At every clock cycle, the control unit produces a control word. Control words are produced in a sequence depending on the type of elliptic curve operation being done. The control word signals control the flow of data and also decide the operations performed on the data. There are 33 control signals (CW0–CW32) that are generated by the control unit. The signals CW0–CW9 control the inputs to the finite field multiplier and the outputs C_0 and C_1 of the ALU. The control lines CW26–CW29 are used for the select lines to the multiplexer in the quadblock. The remaining control bits are used in the register file to read and write data to the registers.

4 Performance Evaluation

In Table 1, the implementation has a better area-time product compared to Chelton and Benaissa [5] and Rebeiro et al. [6], while the latency TL is almost equal. To compare the three designs, the Chelton and Benaissa [5] work has to be scaled since it is GF.

Table 1 Comparing area \times time requirements

Work	Field (m)	Platform	Slices (S)	SS	TL (m s)	TS (m s)	SS \times TS
[5]	163	XC4V200	16,209	33,120	0.019	0.027	894
[6]	233	XC4V140	20,917	20,917	0.029	0.029	606
This work	233	XC4V140	19,226	19,266	0.027	0.027	519

The area is scaled by a factor of $\left(\frac{233}{m}\right)^2$, since area various in the order of $O(n^2)$. The time is scaled by a factor of $\left(\frac{233}{m}\right)$, since the computation time is linear. In Table 1 scaled slices are given by $SS = S\left(\frac{233}{m}\right)^2$ and scaled latency is given by $TS = TL\left(\frac{233}{m}\right)$.

5 Conclusion

We have designed ECCP that utilizes an improved QITA [7] inverse module and conventional HKM as finite field primitives. The proposed ECCP architecture has better area-time product compared to existing techniques. The design is scalable and suitable for implementation in FPGA platforms.

References

1. Miller, V.S.: Use of elliptic curves in cryptography. *Adv. Cryptol.* **218**, 417–426 (1986)
2. Koblitz, M.: Elliptic curve cryptosystems. *Math. Comput.* **48**, 203–209 (1987)
3. Itoh, T., Tsujii, S.: A fast algorithm for computing multiplicative inverse in $GF(2^m)$ using normal bases. *Inf. Comput.* **78**(3), 171–177 (1988)
4. Kaliski, B.S.: The montgomery inverse and its applications. *IEEE Trans. Comput.* **44**(8), 1064–1065 (1995)
5. Chelton, W.N., Benaissa, M.: Fast elliptic curve cryptography on FPGA. *IEEE Trans. Very Large Scale Integr. Syst.* **16**(2), 198–205 (2008)
6. Rebeiro, C., Roy, S.S., Mukhopadhyay, D.: Revisiting the Itoh-Tsujii inversion algorithm for FPGA platforms. *IEEE Trans. VLSI Syst.* **19**(8), 1508–1512 (2011)
7. Venkatasubramani, V.R., Murali, N., Rajaram, S.: An improved quad Itoh–Tsujii algorithm for FPGAs. *IEICE Electron. Express* **10**(18), 1–6 (2013)

Multi-focus and Multi-exposure based Color Image Fusion for Concealed Weapon Detection

Ekta M. Upadhyay and N.K. Rana

Abstract The detection of weapons hidden underneath a person's clothes is of prime importance for security of the public and also safety of public areas such as railway stations, malls, and airports. The screening procedures done manually give inaccurate results when the object is far away from the security personnel and when in a crowded area. The aim was to use a RGB visual image and a corresponding IR image and make a new algorithm for such an application with the help of fusion techniques. This involves not only detecting the weapon correctly, but also retaining the information of the original visual image as much as possible. Thus, using multi-exposure, multi-focus, and DWT fusion techniques, we propose an algorithm to retain the visual as well as the information regarding the weapon in our final output. We may also add that this will also help strengthen evidences against the 'accused' since he/she can be indisputably recognized with the help of the obtained output.

Keywords Concealed weapon detection · HSV · RGB · LAB · Infrared · Infrared negative · Focal length · Exposure

1 Introduction

Fusion of original image and image from different sensor is to acquire all the information present in the source images and also for better interpretation of the image. This process leads to more accurate data interpretation and utility.

E.M. Upadhyay (✉)
JIT University, Chudela, Rajasthan, India
e-mail: e_upadhyay@rediffmail.com

N.K. Rana
Mumbai University, Mumbai, India
e-mail: ranank@rediffmail.com

There is a problem of focal length due to optical lens in digital camera in which part of the image which is out of focus has less depth of field as compared to other parts of the image. Therefore, the solution is multi-focus image fusion [1, 2] which combines several captured images with different focal lengths into one image containing all information. There will be areas in an image of a scene which has high or low exposure and also contains less information than the same area when well exposed. It is essential to capture images at different exposures and then combine them to capture details about an entire scene as proposed in [3, 4]. This fusion process is not only helpful for scene understanding by humans but also by computer vision systems.

Concealed weapon detection (CWD) seems to be a crucial application for dealing with security. The current sensor technologies used for CWD [5, 6] applications are visual, IR, and MMW. Thus, there is an urgent need for a new concealed weapons detection system which detects weapons from a distance in a quick and efficient way. But at the same time, when distance comes into picture, then there is a problem of focal length and exposure. This is the focus of our current efforts, and we have dealt with fusion of visual and IR images taking multi-focus, multi-exposure, multi-view, and multi-modality into consideration.

2 Methodology

The focus of this paper is to obtain a fused image that provides a detailed description of the people in the scene and any hidden weapons detected by the IR image using the best multi-focused [7], multi-exposed visual [8, 9], and IR images and thus obtain the color image with successful CWD.

2.1 Setup for Capturing 360° Images

The setup for capturing the images for CWD is shown in Fig. 1. We have captured the images for front, left, right, and back view of a scene so that we can detect a hidden object in 360°. For each view say front view, we have considered two viewpoints v_1 and v_2 . There are two different focal points f_1 and f_2 for each viewpoint so that we can take care of multi-focus images of the same scene. We get the best focused image from f_1 and f_2 which has best exposed visual and infrared images. Three multi-exposure visuals and infrared images are captured for each focal point of the same viewpoint. We take multi-exposure fusion of visual and infrared images separately. Similarly, the procedure is repeated for the left, right, and back views of a scene.

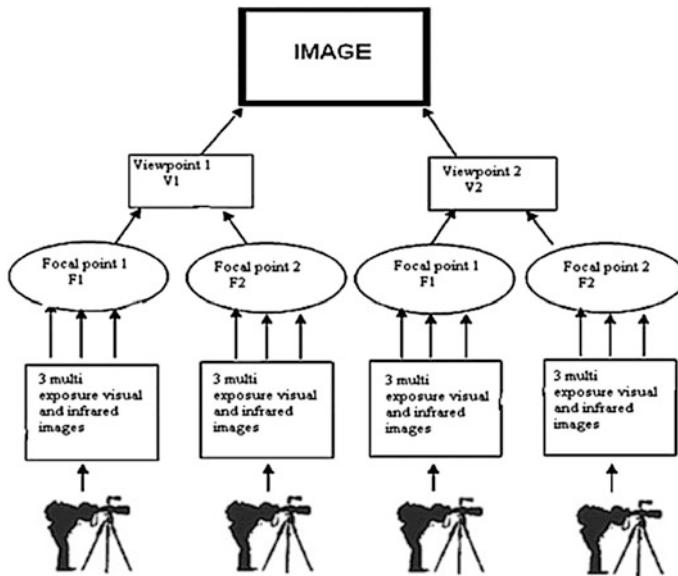


Fig. 1 Setup for capturing images for CWD

2.2 Fusion Methodology

We are provided with multi-exposure visual and IR images for each focal point for every viewpoint of front, right, left, and back view of a scene. We get the best exposed visual and infrared image using multi-exposure fusion. By using multi-focus fusion, we get the best focused image of the best exposed visual and infrared images. The two images are then fused together using DWT fusion [10, 11] to obtain the final fused image taking exposure and focus into consideration in which the weapon is clearly visible.

The algorithm for color image fusion for CWD is as follows:

2.2.1 Input Images

Select the folder name which contains the different multi-exposed images for visual, IR, and IR_{neg} for different focal points (f_1, f_2) of viewpoint1 v_1 .

2.2.2 Apply Multi-exposure Fusion

To obtain the best multi-exposed visual, IR, and IR_{neg} images for each of the focal points, we apply multi-exposure fusion to visual, IR, IR_{neg} images [12–15].

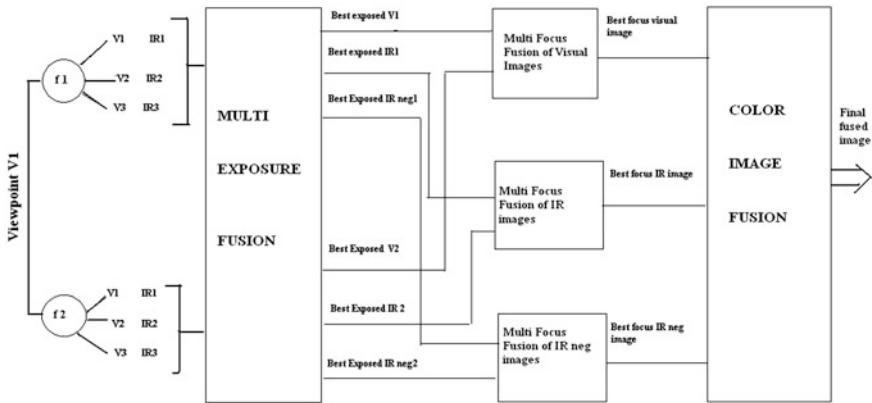


Fig. 2 Block diagram of image fusion methodology for CWD

2.2.3 Apply Multi-focus Fusion

Now, we apply multi-focus fusion to the best exposed visual, IR, and IR_{neg} to obtain the best focus visual, IR, and IR_{neg} images, respectively. The final fusion of visual and infrared image is using DWT. We then pass this as follows:

best focus visual \rightarrow RGB image

best focus of IR image \rightarrow IR image

best focus IR_{neg} \rightarrow IR^{-1}

The block diagram for CWD fusion methodology is shown in Fig. 2.

3 Color Image Fusion Method

To retain the maximum of color information possible, different color models such as HSV [16] and LAB [17] which are better for interpretation have been used. The resultant image will have the best resolution of the visual image, result of concealed weapon detected with the help of IR image. The final fused image for CWD will be obtained as shown in Fig. 3. Firstly, conversion of input RGB image is done into HSV image. HSV color space specifies color information more precisely in terms of brightness and color content. The intensity of the image is represented using V channel of the HSV model which is used in fusion process. The H channel and the S channel will be used to give the fused image a natural color which is similar to the color of the visual image. Besides the HSV color space, the LAB color space is also used.

- Fusion of V_V component in HSV color space output and IR image and also with the reverse polarity IR^{-1} image is conducted using DWT approach.

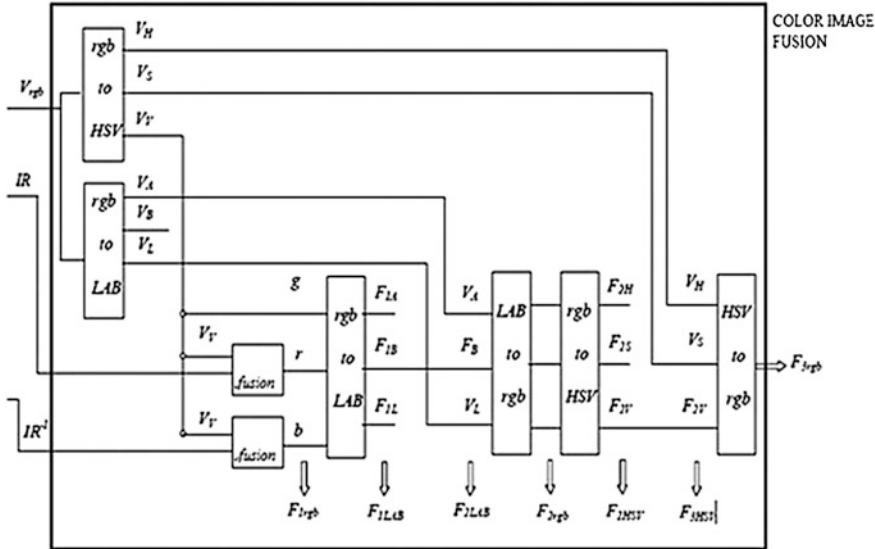


Fig. 3 Block diagram of color image fusion for CWD

- RGB image = R: $V_V + IR$; G: V_V ; B: $V_V + IR^{-1}$. The resulting color image is called a pseudocolor image ($F_{1\text{rgb}}$).
- To retain more color information as visual image, the RGB image is converted to LAB color model denoted as $F_{1\text{LAB}}$ (F_{1L} , F_{1A} , F_{1B}) and V_{LAB} (V_L , V_A , V_B), respectively.
- Then, a new image $F_{2\text{LAB}}$ (F_{2L} , F_{2A} , F_{2B}) is obtained using the following method:

$$(F_{2L}, F_{2A}, F_{2B}) = (V_L, V_A, F_{1B}) \quad (1)$$

- The same equation above is used to convert the LAB model to RGB image.
- Since LAB would retain the color, the brightness and purity of color value needs to be adjusted still, and therefore, the RGB image is converted into HSV color space ($F_{2\text{HSV}}$ (F_{2H} , F_{2S} , F_{2V})); then, a new image $F_{3\text{HSV}}$ (F_{3H} , F_{3S} , F_{3V}) is obtained by carrying out the following procedure:

$$(F_{3H}, F_{3S}, F_{3V}) = (V_H, V_S, F_{2V}) \quad (2)$$

That is, the H channel and S channel of $F_{2\text{HSV}}$ are replaced by the H channel and S channel of V_{HSV} , respectively. Then, we obtain the final RGB image using HSV to RGB conversion technique.

4 Results and Discussion

In order to illustrate the efficiency of the proposed color image fusion method for our CWD application, a number of images are captured using the setup as mentioned in Sect. 2. The corresponding fusion results are presented here which are shown in Figs. 4, 5, 6, 7, 8, 9, 10, 11 respectively. These test images were selected to test the fusion algorithm under various conditions including different view positions for the person, different focal lengths, and different exposures. For all the cases, the visual and IR images have been registered. The input images captured for viewpoint 1 v_1 of front view taking focal length f_1 are as shown in Fig. 4a–c.

The output of best exposed visual, IR, and IR_{neg} images for f_1 and f_2 using multi-exposure fusion is represented in Fig. 6a, b, respectively. Figure 7 is the output for multi-focus fusion on the best exposed output obtained in Fig. 6.

Figure 7 shows the fused image obtained from fusing the gray-level visual image V_V and the reverse polarity IR image using a DWT multi-scale representation. The background is clearer, but the human body is unclear due to its low intensity in the reverse polarity IR image. This fused image is being mapped into the blue channel of pseudocolor image F_{1rgb} . Figure 8 shows the pseudo color fused image F_{1rgb} (red channel: F_{11} , green channel: V_V , and blue channel: F_{12}). The color retained from Fig. 7 is not appropriate.

Figure 9 shows the color image after adjustment in LAB color space (F_{2rgb}).

The color of this image is closer to the color of the visual image than F_{1rgb} . Figure 10 shows the V channel of the image F_{2HSV} , which is obtained by converting the fused color image F_{2rgb} .

The final color image after additional adjustment in the HSV color space F_{3rgb} is shown in Fig. 11. The color of F_{3rgb} is very close to the color of visual image shown in Fig. 4, and F_{3rgb} provides a good depiction of the concealed weapon. It is obvious that this image is better than either of the gray-level fused images.

Figure 4 represents an image where the person is wearing a light color shirt with an object hidden under his clothing. The position of the concealed object can be found in Fig. 11. An advantage of the algorithm is that no training or parameter adjustments are needed in the whole process. The demonstration shown above is for detecting the concealed weapon under a human's clothes for the front view. Similarly, the procedure for detecting concealed weapon is carried out is repeated for left, right, and back view of a scene. The proposed algorithm can also be applied for detecting a concealed weapon hidden under clothes or in bags without a human present.

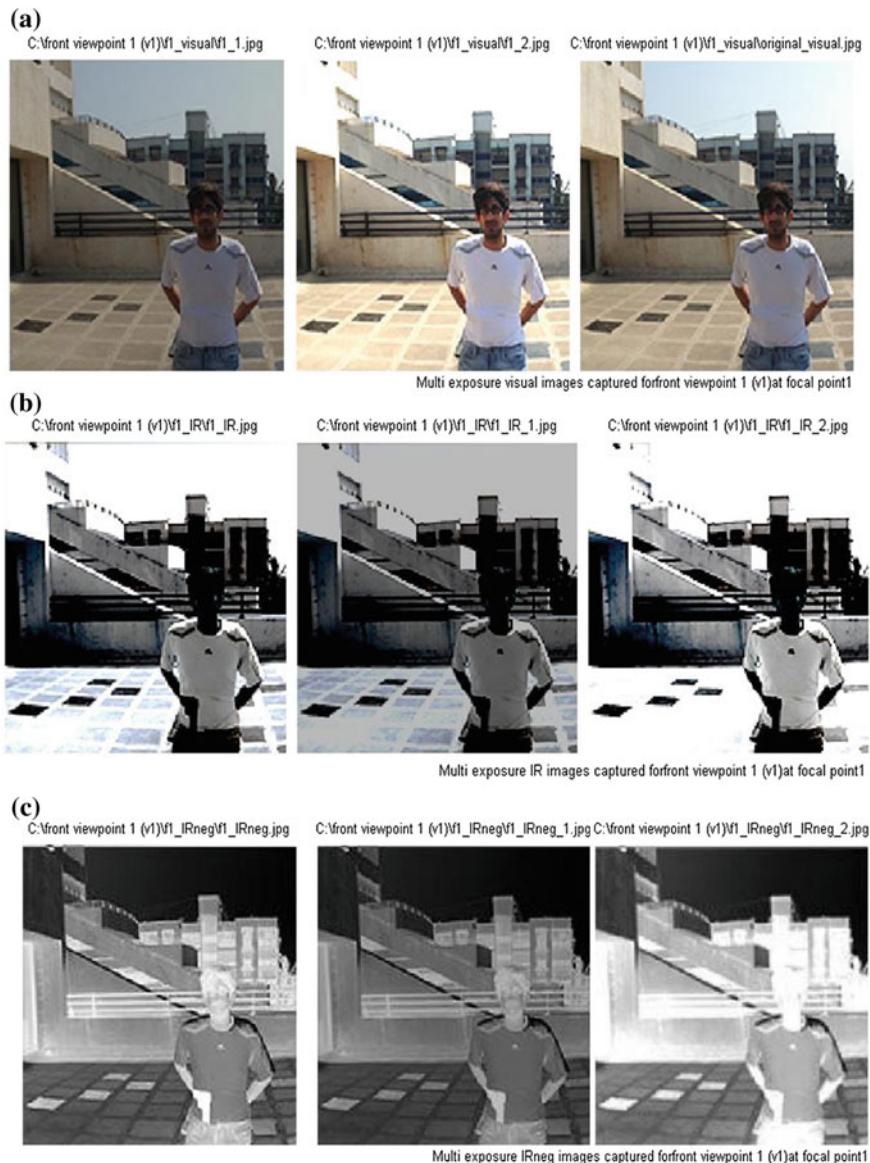


Fig. 4 **a** Multi-exposure visual images for v_1 captured at f_1 . **b** Multi-exposure IR images for v_1 captured at f_1 . **c** Multi-exposure IR_{neg} images for v_1 captured at f_1

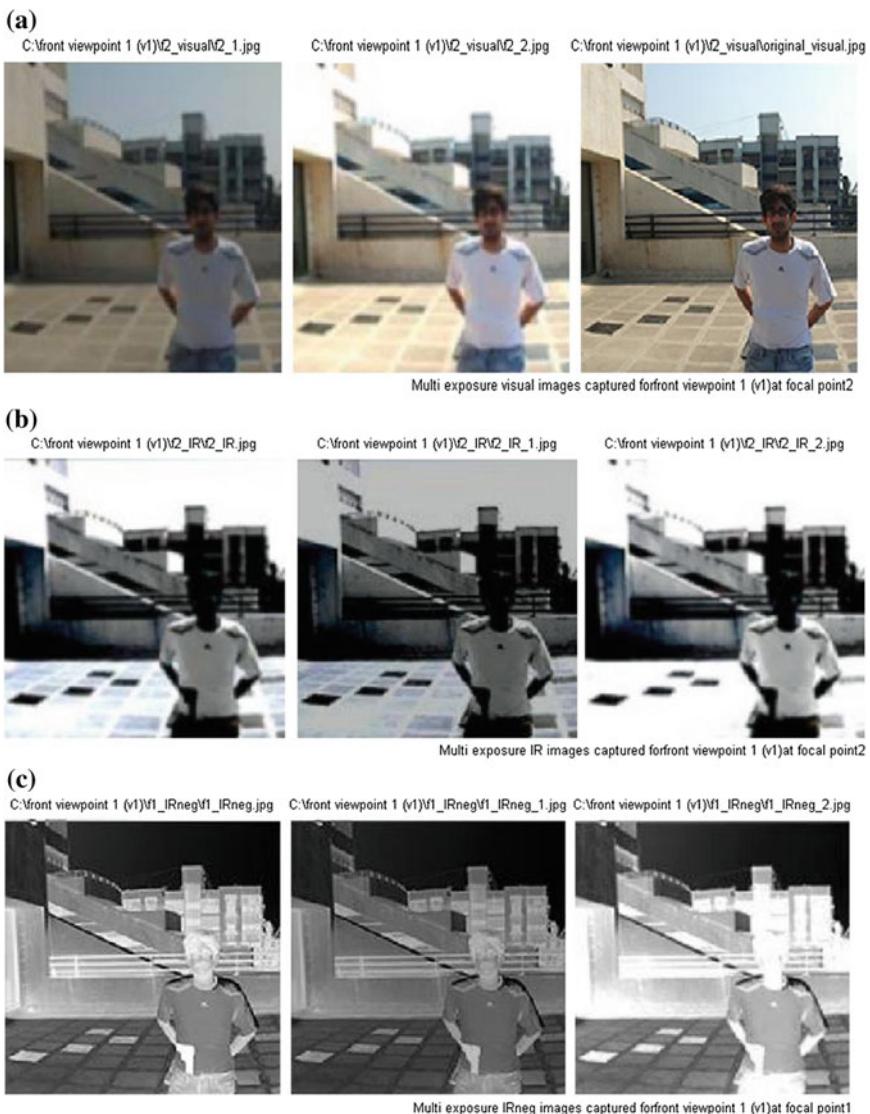


Fig. 5 **a** Multi-exposure visual images for v_1 captured at f_2 . **b** Multi-exposure IR images for v_1 captured at f_2 . **c** Multi-exposure IR_{neg} images for v_1 captured at f_2

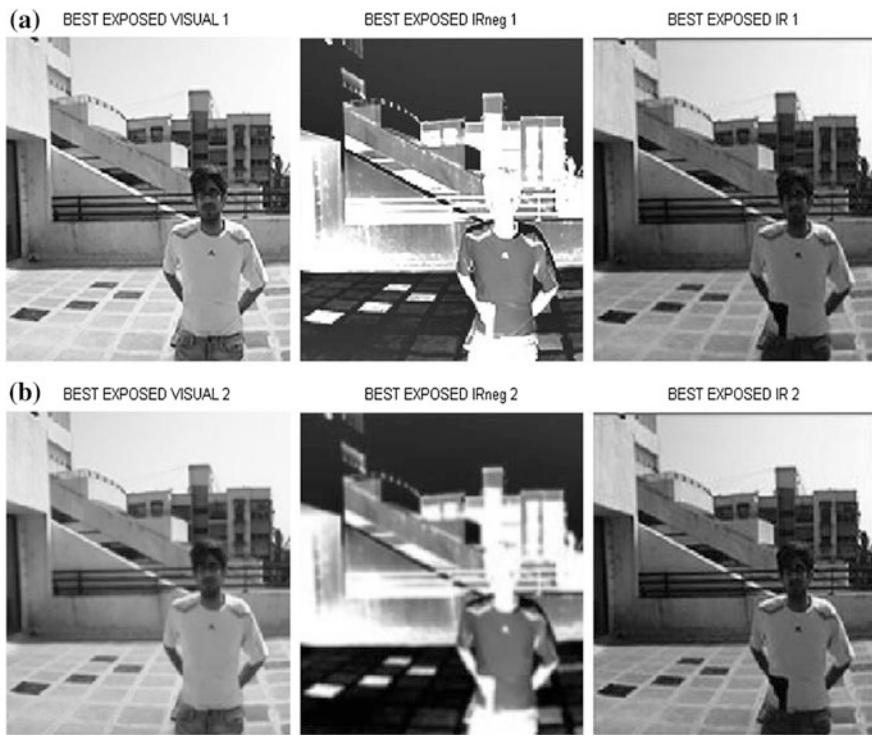


Fig. 6 **a** Multi-exposure fusion results on f_1 images. **b** Multi-exposure fusion results on f_2



Fig. 7 Multi-focus fusion results on best exposed f_1 and f_2 images



Fig. 8 Color fused image $F_{1\text{rgb}}$

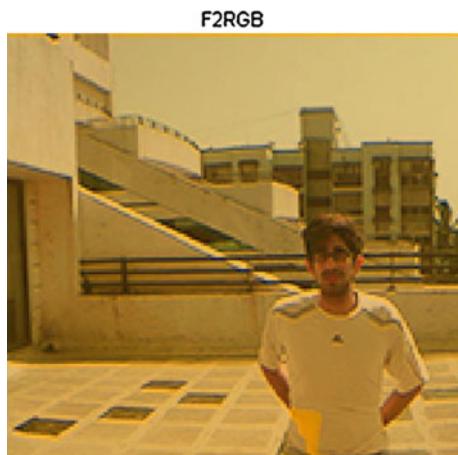


Fig. 9 Color fused image $F_{2\text{rgb}}$

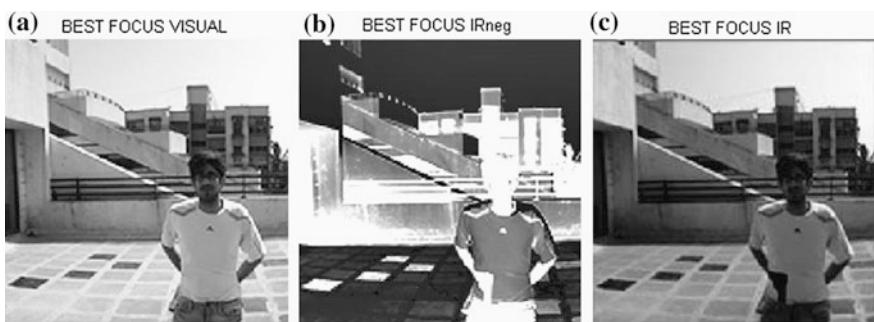
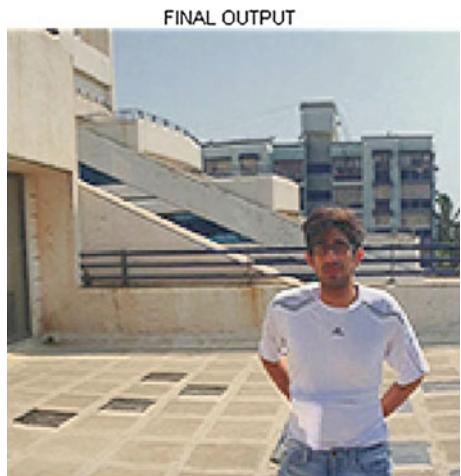


Fig. 10 **a** H channel. **b** S channel. **c** V channel of $F_{2\text{rgb}}$

Fig. 11 Final color fused image $F_{3\text{rgb}}$ after adjustment in $F_{2\text{HSV}}$



5 Conclusion

The aim of the proposed CWD technique is to build a fusion approach for color visual image and IR image of a scene taking multi-view, multi-focus, and multi-exposure into consideration. It is important to detect the hidden weapon accurately, but at the same time, it is also important to recognize the person who is carrying it which is possible only when the color information of the original captured image is retained as much as possible. This experiment would be of interest to researchers in the area of target recognition and also the work could be extended to finding whether the detected object is a weapon or not.

References

1. Hall, D., Llans, J.: Hand book of multisensor data fusion. CRC Press LLC, Boca Raton (2001)
2. Zaveri, T., Zaveri, M.: A novel region based multifocus image fusion method. In: IEEE International Conference on Digital Image Processing, pp. 50–55. Bangkok, Thailand (2009)
3. Wang, C., Ye, Z.: Brightness preserving histogram equalization with maximum entropy: a variational perspective. *IEEE Trans. Consum. Electron.* **51**(4), 1326–1334 (2005)
4. Delac, K., Grgic, M., Kos, T.: Sub-image homomorphic filtering technique for improving facial identification under difficult illumination conditions. In: International Conference on Systems, Signals and Image Processing, (IWSSIP'06), pp. 95–98. Budapest, Hungary, 21–23 Sept 2006
5. Rauscher, D.R., Hartnett, M.P.: Concealed weapon detection program. Decision-Science Applications, Inc., New York (1998)
6. Xue, Z., Blum, R.S., Li, Y.: Fusion of visual and IR images for concealed weapon detection. In: Proceedings of the Fifth International Conference on Information Fusion, vol. 2, pp. 1198–1205 (2002)

7. Upadhyay, E.M., Rana, N.K.: Multi focus fusion based on complex wavelet for surveillance application. *IJETCAS* 13-104, ISSN 2279-0047 (2013)
8. Ardesir Goshtasby, A.: Fusion of multi-exposure images. *Image Vis. Comput.* **23**, 611–618 (2005)
9. Mertens, T., Kautz, J., Van Reeth, F.: Exposure fusion. In: *Proceedings of Pacific Graphics*, Maui, Hawaii (2007)
10. Piella, G.: A general framework for multiresolution image fusion: from pixels to regions. *Inf. Fusion* **4**(4), 259–280 (2003)
11. Mallat, S.G.: A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **11**(7), 674–693 (1989)
12. Upadhyay, E.M., Rana, N.K.: Homomorphic filter based exposure fusion. In: *Proceedings of Fourth International Conference on Recent Trends in Information, Telecommunication and Computing*, ITC 2013. Elsevier, Amsterdam
13. Upadhyay, E.M., Rana, N.K.: Contrast enhancement for multi exposure infrared images. In: *IEEE International Conference on Signal Processing and Integrated Networks*, SPIN 2014. IEEE XPLORE, New York
14. Upadhyay, E.M., Rana, N.K.: Exposure fusion for concealed weapon detection. In: *IEEE 2nd International Conference on Devices, Circuits and Systems—ICDCS 2014*. IEEE XPLORE, New York
15. Upadhyay, E.M., Rana, N.K.: Contrast enhancement in exposure fusion for concealed weapon detection. In: *3rd International Conference on Recent Trends in Engineering and Technology*, ICRTET 2014. Elsevier, Amsterdam
16. Waxman, A.M., Aguilar, M., Baxter, R.A., Fay, D.A., Ireland, D.B., Racamato, J.P., Ross, W. D.: Opponent-color fusion of multi-sensor imagery: visible, IR and SAR. *Proc. IRIS Passive Sens.* **1**, 43–61 (1998)
17. Aguilar, M., Fay, D.A., Ross, W.D., Waxman, A.M., Ireland, D.B., Racamato, J.P.: Real-time fusion of low-light CCD and uncooled IR imagery for color night vision. *Proc. SPIE* **3364**, 124–35 (1998)

A Compact Analytical Model for 2D Triple Material Surrounding Gate Nanowire Tunnel Field Effect Transistors

**D. Saraswathi, N.B. Balamurugan, G. Lakshmi Priya
and S. Manikandan**

Abstract In this paper, we have formulated a two-dimensional (2D) analytical model for surface potential and electric field of a p-type tri-material gate (TMG) gate-all-around (GAA) nanowire tunneling field effect transistor (TFET). The proposed model includes the effect of voltage at the drain end, work functions of the gate metal, oxide thickness, and radius of the silicon nanowire. Incorporating the advantages of surrounding gate with three different gate material work functions, this nanostructure of TMG GAA TFET is proposed and shows improved electrical characteristics. Also a step change of potential along the channel is obtained, which screens the region R1 from the drain potential variations, thus reducing the drain control over the channel. This results in appreciable diminishing of short channel effects and hot carrier effects, which are investigated by developing a 2D analytical model. The analytical results are compared and verified with the simulation results.

Keywords Triple material gate (TMG) · Gate-all-around (GAA) · Tunneling field effect transistor (TFET) · Nanowire · Two-dimensional (2D) modeling · Surface potential · Electric field

D. Saraswathi (✉) · N.B. Balamurugan · G.L. Priya · S. Manikandan
Department of Electronics and Communication, Thiagarajar College of Engineering,
Madurai, India
e-mail: saraswathi78ece@gmail.com

N.B. Balamurugan
e-mail: nbbalamurugan@tce.edu

G.L. Priya
e-mail: priya0217@gmail.com

S. Manikandan
e-mail: maniaugust36@gmail.com

1 Introduction

Studies on novel device structures for VLSI applications are being done extensively these days to provide alternatives to conventional CMOS transistors. This is because MOSFETs scaled to lengths below 100 nm face several problems with regard to leakage currents in the OFF-state, subthreshold swing (SS), drain induced barrier lowering (DIBL), and numerous other short channel effects (SCE). An attractive alternative is tunneling field effect transistor (TFET) [1–4], which has SS below 60 mV/decade, low OFF-state leakage currents, and diminished short channel effects (SCE). However, since the source of carriers is band-to-band tunneling, TFETs have a low ON-state current I_{ON} and it does not meet ITRS requirements. Also, a gate-all-around (GAA) structure provides an improved SS and a solution to SCE and DIBL because of the better electrostatic control over the channel. Further, a higher I_{ON} per unit device area is achieved due to the device geometry. However, the problems of delayed saturation, low I_{ON} , and leakage currents remain in a GAA structure. Thus, to optimize the GAA structure, a triple material gate (TMG) structure is proposed. A TMG GAA nanowire TFET has a lower work function for the gate at the source side, and it increases as it moves toward the drain for an n-channel TFET and vice versa for a p-channel TFET. The TMG GAA nanowire TFET will provide a higher I_{ON} due to the increased tunneling by a metal of lower work function. Thus, we will get a better I_{ON}/I_{OFF} ratio and a better SS. Thus, a TMG GAA nanowire TFET provides a higher I_{ON} with reduced SCEs and better switching characteristics as compared to a planar TFET. Although TFETs using other materials are being studied, there is a great interest in silicon TFETs with improved I_{ON} because of their suitability in silicon-based CMOS technology [5–8]. Several models have been developed for conventional TFETs but a model for TMG GAA TFET is yet to be developed. The objective of this work is, therefore, to develop a 2D analytical model for the TMG GAA nanowire TFET. In this work, we begin with the modeling of the surface potential along the channel using a 2D model for solving the Poisson's equation in the silicon channel. We then find the average electric field in the tunneling region and derive the drain current using Kane's model. The accuracy of our models is validated by comparing the results given by our models with three-dimensional numerical simulations.

2 Proposed Structure for Tri-material Nanowire Tunnel FET

Figure 1 represents the cross-sectional view of tri-material TFET with $L = 90$ nm as the length of the gate metal M. The gate metal M is made of three materials M_1 , M_2 , and M_3 of different work functions deposited over respective lengths L_1 , L_2 , and L_3 on the gate oxide layers. The device is composed of gate material 1 with the work function $\phi_{M_1} = 4.8$ eV (Au), gate material 2 with $\phi_{M_2} = 4.6$ eV (Mo), and gate material 3 with $\phi_{M_3} = 4.4$ eV (Ti).

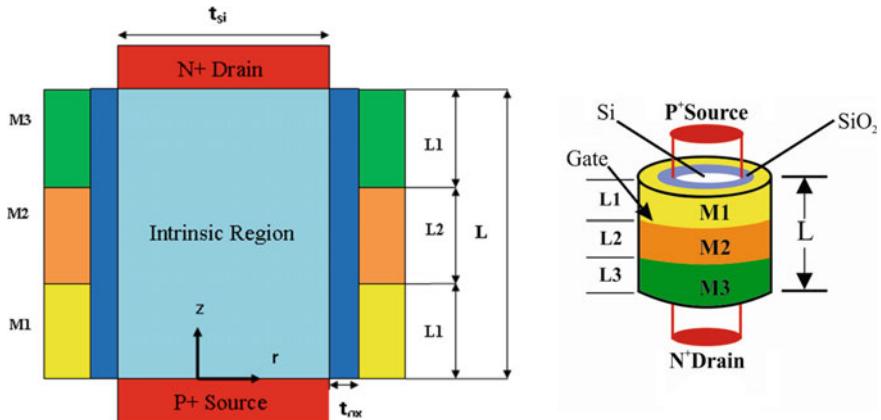


Fig. 1 Cross-sectional view of tri-material nanowire tunnel FET and its structure

2.1 Model Derivation

The 2D Poisson's equation for potential $\phi(r, z)$ in fully depleted TMFET's is given as follows:

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \phi(r, z)}{\partial r} \right) + \frac{\partial^2 \phi(r, z)}{\partial z^2} = \frac{qN_a}{\epsilon_{si}} \quad (1)$$

where $\phi(r, z)$ is the electrostatic potential in the silicon film, q is the electronic charge, and ϵ_{si} is the permittivity of the silicon film. The potential profile in the vertical direction can be approximated by simple parabolic function as,

$$\phi(r, z) = s_1(z) + s_2(z)r + s_3(z)r^2 \quad (2)$$

The Poisson's equation is solved separately using boundary conditions.

- (a) The surface potential $\phi_s(z)$ is a function of z only.

$$\psi(T_{Si}, z) = \psi_s(z) \quad (3)$$

- (b) The electric field in the center of the silicon pillar is zero.

$$\left. \frac{\partial}{\partial r} \psi(r, z) \right|_{r=0} = 0 \quad (4)$$

(c) The electric field at the silicon/oxide interface is continuous.

$$\left. \frac{\partial \psi(r, z)}{\partial r} \right|_{r=T_{\text{Si}}} = \frac{-C_{\text{ox}}}{\epsilon_{\text{Si}}} (\psi_G - \psi_S(z)) \quad (5)$$

where

$$C_{\text{ox}} = \frac{\epsilon_{\text{ox}}}{T_{\text{Si}} \ln \left(1 + \frac{T_{\text{ox}}}{T_{\text{Si}}} \right)}$$

where

$$\psi_{G_i} = V_{GS} - \phi_{m_i} + \chi + \frac{E_g}{2}$$

The surface potential can be written as

$$\frac{\partial^2 \psi_S}{\partial Z^2} - \frac{2C_{\text{ox}}}{T_{\text{Si}} \epsilon_{\text{Si}}} \psi_S = \frac{qN_S}{\epsilon_{\text{Si}}} - \frac{2C_{\text{ox}}}{T_{\text{Si}} \epsilon_{\text{Si}}} \psi_S.$$

2.2 Surface Potential

Thus, the surface potential for the three regions is given by a general form as,

$$\psi_{S_i}(Z) = C_i \exp \left(\frac{Z - L_i}{L_d} \right) + D_i \exp \left(\frac{Z - L_i}{L_d} \right) + \psi_{G_i} - \frac{qN_i L_d^2}{\epsilon_{\text{Si}}} \quad (6)$$

where

$$L_d^2 = \frac{T_{\text{Si}}^2 \ln \left(1 + \frac{T_{\text{ox}}}{T_{\text{Si}}} \right) \epsilon_{\text{Si}}}{2 \epsilon_{\text{ox}}}$$

The boundary conditions given below are used to determine the constants C_i and D_i values of the surface potential equation for the three regions.

(a) Surface potential at source is equal to V_{bi} :

$$\psi_{S1}(0) = V_{bi} \quad (7)$$

(b) At the boundary of region, R1 is equal to:

$$\psi_{S1}(L_1) = \psi_{C1} \quad (8)$$

(c) Electric field at the boundary of region R1 is equal to zero:

$$\frac{\partial \psi_{S1}(L_1)}{\partial Z} = 0 \quad (9)$$

(d) Surface potential at other regions is given by:

$$\psi_{S2}(-L_2) = \psi_{Ct} \quad (10)$$

$$\psi_{S3}(L_3) = \psi_{Ca} \quad (11)$$

(e) The electric field is given by:

$$\frac{\partial \psi_{S2}(-L_2)}{\partial Z} = 0 \quad (12)$$

$$\frac{\partial \psi_{S3}(L_3)}{\partial Z} = 0 \quad (13)$$

Thus, the value of constants obtained for different regions are given as follows:

$$C_1 = D_1 = \left(V_{bi} - \psi_{Gt} + \frac{qN_S L_d^2}{\varepsilon_{Si}} \right) / \left(2 \cosh \left(\frac{-L_t + L_1}{L_d} \right) \right) \quad (14)$$

$$C_2 = D_2 = \frac{\psi_{C1} - \psi_G + \frac{qN_S L_d^2}{\varepsilon_{Si}}}{2} \quad (15)$$

$$C_3 = D_3 = \frac{\psi_{C3} - \psi_G + \frac{q(N_S + n)L_d^2}{\varepsilon_{Si}}}{2} \quad (16)$$

The slope of the surface potential decreases along the channel length and hence the shortest tunneling length L_T must lie between the source and the point where the potential falls by E_G/q and can be written as follows:

$$L_T = Z(\psi_{Source}) - Z(\psi_{Source} - E_g/q) \quad (17)$$

$$Z(\psi_S) = L_d \cosh^{-1}((\psi_S - \psi_G)/(\psi_C - \psi_G)) \quad (18)$$

Using L_T , the average electric field along the tunneling path F_{avg} can be calculated as follows:

$$F_{\text{avg}} = E_G/qL_T \quad (19)$$

Substituting F_{avg} into the Kane's model for band-to-band tunneling, we get the expression for the drain current as follows:

$$I_D = A_K \times L_T \times F_{\text{avg}}^2 \times \exp(-B_K/F_{\text{avg}}). \quad (20)$$

3 Results and Discussions

Surface potential and electric field for TMG GAA Nanowire TFET are calculated and plotted for various device parameters. The proposed analytical model results are compared with DMG TFETs. The obtained results are also validated with MEDICI simulation model results.

Figure 2 shows the variation of surface potential along the length of the channel for various analytical models. The results of the proposed work are compared with SMG GAA nanowire TFET and DMG GAA nanowire TFET. It is observed that TMG GAA nanowire TFET has a better potential than the other two models. The results of the proposed work are also verified with MEDICI simulation model.

Figure 3 shows the variation of lateral electric field along the channel length for different values of V_{ds} . It is clearly observed that there is a peak electric field toward the drain side, which results in larger I_{OFF} current in the weak inversion region. It is also inferred that lateral electric field variation is less dominant in the TMG GAA

Fig. 2 Surface potential variation along the channel for various models

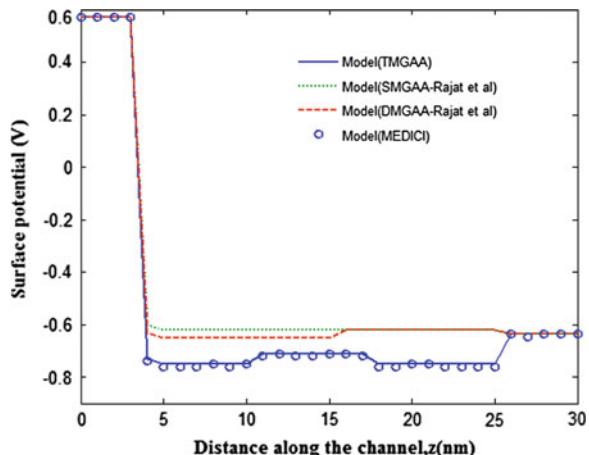
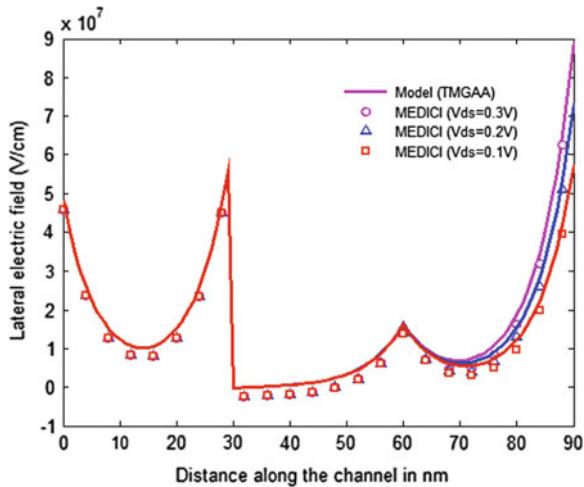


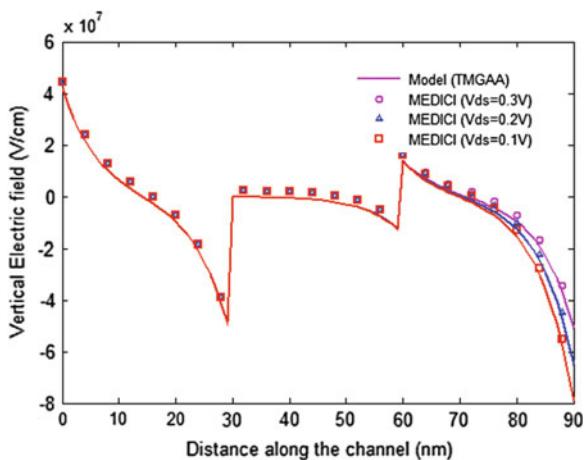
Fig. 3 Variation of lateral electric field along the channel length for different values of V_{ds}



nanowire TFET, because even when V_{ds} is varied the variation of lateral electric field is less when compared to vertical electric field.

Figure 4 shows the vertical electric field of TMG GAA nanowire TFET along the channel length for various values of drain to source biases. As the material work function toward the drain side (i.e., region under gate material 3) is low when compared to the source side, the peak electric field at drain side is very low. The tunneling gate is completely screened from the drain potential variations due to higher material work function. This minimizes the short channel effects such as DIBL and results in a subthreshold slope to be less than 50 mV/dec. The simulation results matches well with the proposed analytical model results.

Fig. 4 Vertical electric field of TMG GAA nanowire TFET along the channel length for different values of V_{ds}



4 Conclusion

The effectiveness of the triple material gate concept to the nanowire TFET has been examined for the first time by developing a 2D analytical model. Expressions for the surface potential, lateral electric field, and vertical electric field have been obtained, proving the superiority of TMSG TFET in suppressing the SCEs over the conventional DMSG structure. The results obtained from the simulation agree well with the theoretical results. It is apparent from the results that the TMSG architecture exhibits improved gate controllability over the channel. Results also emphasize that the device optimization in terms of the metal gate length ratios, work function differences, and barrier layer thicknesses leads to improvement in the device performance. Thus, the introduction of the TMSG structures over their surrounding gate counterparts offers a new way of improving the short channel behavior of nanowire TFETs and also shows the potential for many future applications where high performance TFETs, with gate lengths down to sub-100 nm, are required. The analytical results of the proposed work are compared and verified with the simulation results obtained by MEDICI.

References

1. Rajat, V., Jagadesh Kumar, M.: A pseudo 2D-analytical model of dual material gate all-around nanowire tunneling FET. *IEEE Trans. Electr. Devices* (2014)
2. Shao, Q., Zhao, C., Wu, C., Zhang, J., Zhang, L., Yu, Z.: Compact model and projection of silicon nanowire tunneling transistors (NWTFTs). In: International Conference of Electronic Devices and Solid-State Circuits (EDSSC) (2013)
3. Saurabh, S., Kumar, M.J.: Estimation and compensation of process induced variations in nanoscale tunnel field effect transistors (TFETs) for improved reliability. *IEEE Trans. Device Mater. Reliab.* **10**(3), 390–395 (2010)
4. Verhulst, A.S., Sorée, B., Leonelli, D., Vandenberghe, W.G., Groeseneken, G.: Modeling the single-gate, double-gate, and gate-all-around tunnel field-effect transistor. *J. App. Phys.* **107**(2), 024518 (2010)
5. Shen, C., Ong, S.-L., Heng, C.-H., Samudra, G., Yeo, Y.-C.: A variational approach to the two-dimensional nonlinear Poisson's equation for the modeling of tunneling transistors. *IEEE Electr. Device Lett.* (2008)
6. Santra, AA.: A subthreshold analysis of triple-material cylindrical gate-all-around (TM-CGAA) MOSFETs. National Institute of Technology, Rourkela (2013)
7. Kumar, S.P., Agrawal, A., Chaujar, R., Kabra, S., Gupta, M., Gupta, R.S.: Threshold voltage model for small geometry AlGaN/GaN HEMTs based on analytical solution of 3-D Poisson's equation. *Microelectron. J.* **38**(101), 1013–1020 (2007)
8. Chaudhry, A., Kumar, M.J.: Controlling short-channel effect in deep-submicron SOI MOSFETs for improved reliability: a review. *IEEE Trans. Electr. Devices* **4**(1), 99–109 (2006)

Application of Particle Swarm Optimization Technique in Hexagonal and Concentric Hexagonal Antenna Array for Side Lobe Level Reduction

Rajesh Bera, Durbadal Mandal, Rajib Kar and Sakti Prasad Ghoshal

Abstract In this paper, particle swarm optimization (PSO) method, which represents a recent approach for optimization problems in electromagnetic, is applied for array pattern synthesis of hexagonal array (HA) and concentric hexagonal array (CHA) of uniformly excited isotropic antennas which can generate directive beam with minimum relative side lobe level (SLL). Two examples has been presented and solved. In first example, the PSO is used to determine an optimal set of ‘ON-OFF’ elements in a 12-element thinned HA and 24-element thinned CHA, and in second example, PSO is used to determine an optimal set of amplitude distributions in a 12-element HA and a 24-element CHA that provide a radiation pattern with maximum SLL reduction. Optimization is done without significant change in the value of first null beam width (FNBW). Simulation results show that, in first example, the number of effective antenna elements can be brought down from 12 to 6 totals with simultaneous reduction in SLL by -23.85 dB for HA and from 24 to 11 totals with simultaneous reduction in SLL by -20.71 dB for CHA. In second example, SLL is reduced to less than -31 and -27 dB for HA and CHA, respectively, by optimizing inter-element spacing and inter-ring spacing along with amplitude using PSO.

Keywords Thinned array · Particle swarm optimization · Hexagonal array · Concentric hexagonal array · Side lobe level

R. Bera (✉) · D. Mandal · R. Kar

Department of Electronics and Communication Engineering,

National Institute of Technology Durgapur, Durgapur, West Bengal, India

e-mail: rajeshkiit12@gmail.com

D. Mandal

e-mail: durbadal.bittu@gmail.com

R. Kar

e-mail: rajibkarece@gmail.com

S.P. Ghoshal

Department of Electrical Engineering, National Institute of Technology Durgapur, Durgapur, West Bengal, India

e-mail: spghoshalnitdgp@gmail.com

1 Introduction

Usually, the radiation pattern of a single antenna element is relatively wide and each element provides low directivity. Antenna arrays increase the directivity without enlarging the size of single element. Generally, the overall array properties such as directivity and gain, direction of maximum directivity, first null beam width (FNBW), side lobe level (SLL), and half-power beam width (HPBW) can be controlled and optimized by adjusting the number of elements, the inter-element spacing, their excitation coefficients, their relative phases, the geometrical arrangement of the overall array (linear, circular, elliptical, etc.), and the relative pattern of the individual elements [1].

Nowadays, one of the most recent innovations for wireless communications to overcome the problem of increasing demand for capacity is to deploy smart antenna [2], which can be realized as the antenna directing a beam toward the communication partner only. Smart antennas not only increase the capacity, they have also potential to introduce new services, increased range, more security, reduced multipath propagation, etc. Narrower main beam and more nulls in the pattern can resolve the signals of interest (SOI) more accurately and allow the smart antenna system to throw out more signals not of interest (SNOI).

A linear array has excellent directivity and it can form the narrowest main lobe in a given direction, but it does not work well in all azimuth directions. Since a circular array (CA) does not have edge elements, directional patterns synthesized with a CA can be electronically rotated in the surface of the array without a significant change of the beam shape [3]. CA pattern has no nulls in azimuth plane [1]. In smart antenna applications to reject SNOI, the array pattern should have several nulls in the azimuth plane. This can be implemented by the use of elliptical arrays instead of circular arrays [4].

The CA antenna is of high side lobe geometry. If the inter-element distance of array elements is decreased to reduce the side lobes, the mutual coupling influence becomes more significant. For mitigating high SLLs, concentric arrays are utilized in [5]. Also concentric CA antennas have several advantages including the flexibility in array pattern synthesis and design both in narrowband and broadband beam forming applications [6].

Thinning an array means switching off some elements in a uniformly spaced or periodic array (linear or planer) to generate a pattern with low SLL. In this proposed method, the locations of the elements are fixed and all the elements have two states either “on” or “off” (similar to logic “1” and “0” in digital domain), depending on whether the element is connected to the feed network or not. In the “off” state, either the element is passively terminated to a matched load or open circuited. If there is no matching between the elements, it is equivalent to removing “off” element from the array. The advantages of thinned array are the number of antenna elements can be reduced, and the large side lobes in thinned array are easily eliminated where performance of array is not significantly degraded. There are

many published articles [7–9] dealing with the synthesis of thinned array. Element behavior in a thinned array is described in [10].

There are several disadvantages when using classical optimization methods. Some of them are the following: (i) when the number of solution variables, i.e., the size of the solution space, increases, starting point is highly sensitive because the solution of an optimization problem solved using a classical method depends on the arbitrarily chosen initial solution, (ii) continuous and differentiable objective functions are required because discrete variables are difficult to handle using classical methods of optimization, and (iii) a particular classical optimization method may not be appropriate to solve various problems. Thus, it is necessary to develop an efficient and robust optimization method. There are various evolutionary optimization tools for thinning such as genetic algorithm (GA) and particle swarm optimization (PSO) [11–14] which do not suffer from above disadvantages. The PSO algorithm has proven to be a better substitute to other evolutionary algorithms such as GA and ant colony optimization (ACO) for optimal design of antenna array.

2 Design Equations

The total field of an array, according to pattern multiplication principle, can be formed by multiplying the field of a single element at a selected reference point and the array factor. The array factor can be formulated by replacing the actual elements with isotropic (point) sources. The array factor of an array in general can be written as [1]

$$\text{AF}(\theta, \varphi) = \sum_{n=1}^N A_n e^{j(\alpha_n + kR_n) a_r} \quad (1)$$

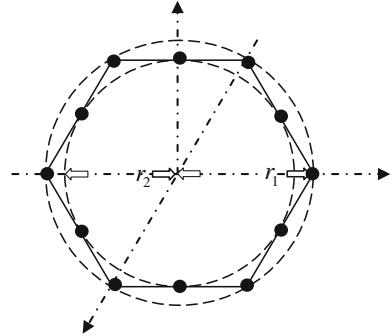
where N is the number of elements, A_n is the relative amplitude of n th element, α_n is the relative phase of n th element, a_n is the position vector of n th element depends on array geometry, a_r is the unit vector of observation point in spherical coordinate, and k is the wave number.

The array factor of an N -element circular array with its center in origin of x – y plane is [3]

$$\text{AF}(\theta, \varphi) = \sum_{n=1}^N A_n e^{jkr \sin \theta (\cos \varphi_n \cos \varphi + \sin \varphi_n \sin \varphi)} \quad (2)$$

where r is radius of the circle, ϕ is the angle between positive section of x axis and the observation point in the space, θ is the angle between positive section of z axis and the observation point in the space, and $2\pi(n - 1)/N$ is the angle in the x – y plane between the x axis and the n th element.

Fig. 1 Hexagonal array (HA) structure



The hexagonal array (HA) in which the peripheral curve of its vertices is a circle can be treated as consisting of two concentric N -element circular arrays of different radii r_1 and r_2 , respectively. Figure 1 gives a simple example of HA with $2N$ elements ($N = 6$), N of which are located at the vertices of the hexagon and the other N elements are placed at the middle of each line of the hexagon, respectively. Using (2) and above explanation, the array factor of HA is

$$\text{AF}(\theta, \varphi) = \sum_{n=1}^{N} [A_n e^{jkr_1 \sin \theta (\cos \varphi_{1n} \cos \varphi + \sin \varphi_{1n} \sin \varphi)} + B_n e^{jkr_2 \sin \theta (\cos \varphi_{2n} \cos \varphi + \sin \varphi_{2n} \sin \varphi)}] \quad (3)$$

$$\begin{aligned} r_2 &= r_1 \cos(\pi/N) \\ r_1 &= d_e / \sin(\pi/N) \end{aligned} \quad (4)$$

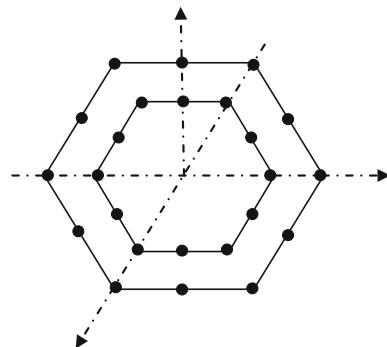
where $\varphi_{1n} = 2\pi(n - 1)/N$ is the angle in the x - y plane between the x axis and the n th element at the vertices of the hexagon, $\varphi_{2n} = \varphi_{1n} + \pi/N$ is the angle in the x - y plane between the x axis and the n th element at the middle of each line of the hexagon, A_n and B_n are the relative amplitudes of n th element placed at the vertices and the middle of the hexagon, respectively, and d_e is the spacing between two elements.

The array factor of concentric hexagonal array (CHA), as shown in Fig. 2, can be found by summation of the array factors of M concentric HA

$$\begin{aligned} \text{AF}(\theta, \varphi) = \sum_{m=1}^M \sum_{n=1}^{N} &[A_{nm} e^{jkr_{1m} \sin \theta (\cos \varphi_{1n} \cos \varphi + \sin \varphi_{1n} \sin \varphi)} \\ &+ B_{nm} e^{jkr_{2m} \sin \theta (\cos \varphi_{2n} \cos \varphi + \sin \varphi_{2n} \sin \varphi)}] \end{aligned} \quad (5)$$

where $2N$ is the number of elements on each hexagon, M is the number of concentric hexagons, A_{nm} and B_{nm} are the amplitudes of excitation currents, r_{1m} is the

Fig. 2 Concentric hexagonal array (CHA) structure



radius of the circle on which elements are placed at the vertices of the m th hexagon, and r_{2m} is the radius of the circle on which elements are placed at the middle of each line of the m th hexagon.

$$\begin{aligned} r_{1m} &= r + (m - 1)d \\ r_{2m} &= r_{1m} \cos(\pi/N) \end{aligned} \quad (6)$$

where r is the radius of smallest circle peripheral to the smallest hexagon which elements lie on its vertices and d is the spacing between hexagons.

3 Particle Swarm Optimization

PSO is a flexible, robust population-based stochastic search or optimization technique with implicit parallelism, which can easily handle with non-differential objective functions, unlike traditional gradient-based optimization methods. PSO is less susceptible to getting trapped on local optima unlike GA and simulated annealing. Eberhart et al. developed PSO concept similar to the behavior of a swarm of birds. PSO is developed through simulation of bird flocking and fish schooling in multidimensional space. Bird flocking optimizes a certain objective function. Each particle knows its best value so far (pBest). This information corresponds to personal experiences of each particle. Moreover, each particle knows the best value so far in the group (gBest) among all pBests. Namely, each particle tries to modify its position using the following information:

- The distance between the current position and the pBest.
- The distance between the current position and the gBest.

After finding the two best values, the particle updates its velocity and positions with following Eqs. (7) and (8) shown below.

$$\begin{aligned} V_{j,i+1} = & (w_{i+1} * V_{j,i} + C_1 * rnd_1 * (p\text{Best}_j - \text{Position}_{j,i}) \\ & + C_2 * rnd_2 * (g\text{Best}_i - \text{Position}_{j,i})) \end{aligned} \quad (7)$$

$$\text{Position}_{j,i+1} = \text{Position}_{j,i} + V_{j,i+1} \quad (8)$$

where $V_{j,i}$ is the velocity of vector j at iteration i ; w is the weighting function; C_1 and C_2 are called social and cognitive constants, respectively; rnd_i is the random number between 0 and 1; $\text{Position}_{j,i}$ is the current position of vector j at iteration i ; $p\text{Best}_j$ is the pBest of vector j ; $g\text{Best}_i$ is the gBest of the group of vectors at iteration i . Normally, $C_1 = C_2 = 1.5-2.05$.

Inertia weight (w_{i+1}) at $(i + 1)$ th cycle is

$$w_{i+1} = w_{\max} - \frac{w_{\max} - w_{\min}}{i_{\max}}(i + 1)$$

where $w_{\max} = 1.0$; $w_{\min} = 0.4$; i_{\max} = maximum number of iteration cycles.

The pseudo-code of the procedure is as follows:

For each particle

 Initialize the particle

End

Do

For each particle

 Calculate the cost value

If the cost value is better than the best cost value (**pBest**) in history

 Set current value as the new **pBest**

End

 Choose the particle with the best cost value of all the particles as the **gBest**

For each particle

 Calculate particle velocity according to Eq. (7)

 Update particle position according to Eq. (8)

End

While maximum iterations or minimum error criteria are not attained, the flowchart of the general PSO algorithm is depicted in Fig. 3.

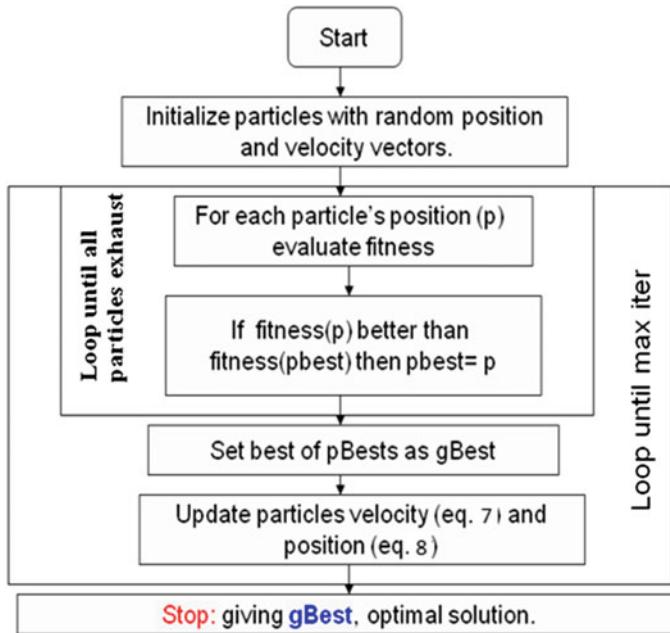


Fig. 3 Flowchart of PSO algorithm

The first and most important parameter in antenna pattern synthesis is the normalized SLL that is desired to be as low as possible. Two different fitness functions have been formulated in Eqs. (9) and (10) for the synthesis of two examples of two different types of amplitude distributions.

First optimization task in the array designing problem is modeled as a minimax problem, i.e., minimization of maximum SLL. Cost function (CF) is formulated to meet the corresponding design goal as follows:

$$\text{CF1} = \min(\max\{20 \log |\text{AF}(\theta)|\}) \quad (9)$$

Subject to $\theta \in \{[-90^\circ, -|\text{FN}|^\circ] \text{ and } [|\text{FN}|^\circ, 90^\circ]\}$

where ‘FN’ is the first null in degree in a particular array pattern.

Second cost function (CF2) is formulated to meet the corresponding designing goal as follows.

$$CF2 = C_1[(SLL_c - SLL_d)] + C_2[(FNBW_c - FNBW_d)] \quad (10)$$

where SLL_d and SLL_c are desired and computed values of SLL, and $FNBW_d$ and $FNBW_c$ are desired and computed values of FNBW, respectively. The first term in (10) is used to reduce the SLL to a desired level. The second term in (10) is introduced to keep FNBW of the optimized pattern to a desired level. C_1 and C_2 are

weighting coefficients to control the relative importance of each term in (10). Because the primary aim is to achieve a minimum SLL, the value of C_1 is higher than the value of C_2 .

The population size using PSO, by which the antenna array is optimized, is 50 and maximum number of iteration cycle is 100. The minimum CF values against number of iteration cycles are recorded to get the convergence profile of the algorithm.

4 Computational Results

The paper describes the application of PSO algorithm for reduction of the maximum SLL without prefixing the value of FNBW and percentage of thinning of HA and two-ring uniform CHA of isotropic elements.

In the analysis of first example, we consider a HA of $2N$ elements ($N = 6$), so the total number of isotropic elements in such an array is 12 and a CHA of two concentric hexagons ($M = 2$) of isotropic elements where on each hexagon $2N$ number of elements ($N = 6$) are there. So in case of CHA, total number of elements we consider is 24.

The inter-element spacing is kept fixed at $d = \lambda/2$ for both of the structures (HA and CHA), and the spacing between two hexagons in CHA is also fixed at $\lambda/2$. Performances of both optimized thinned array antennas with variation of thinning percentage are tabulated in Table 1. The optimal excitation amplitudes for thinned HA and CHA are shown in Table 2.

The percentage of thinning of the array is also calculated. Thinning is the ratio of the number of ‘OFF’ elements to the total number of elements in the array.

For some good cases of different executions, normalized absolute power pattern of optimized thinned HA and CHA are shown in Figs. 4, 5, 6, and 7.

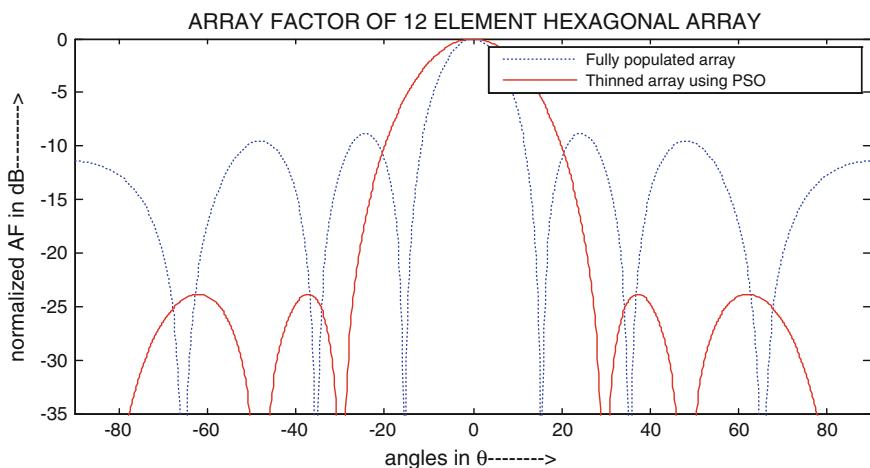
In the second example, a 12-element HA and a 24-element double-ring CHA of isotropic elements are used for synthesis. PSO technique is used to determine an optimal set of non-uniformly excited array that provide a radiation pattern with maximum SLL reduction. In this example, CF2 is used for the synthesis of the array using PSO.

Table 1 Some best performances of thinned HA and CHA

Array structure	Number of ‘ON’ elements		Maximum SLL (dB)		FNBW in degree	Percentage of thinning
	Thinned array	Fully populated array	Thinned array	Fully populated array	Thinned array	Thinned array (%)
HA	06	12	-23.85	-8.87	59.60	50
CHA	11	24	-20.71	-12.10	53.60	54.17
CHA	17	24	-22.96	-12.10	51.20	29.17
CHA	16	24	-20.24	-12.10	47.20	33.33

Table 2 The optimal excitation amplitudes corresponding to the array configuration shown in Table 1

Array structure	Ring number	Placement of elements	Distribution of ON and OFF elements	SLL _{max} (dB)
HA	Ring 1	Vertices	1 1 0 1 0 0	-23.85
		Middle	1 0 1 0 0 1	
CHA	Ring 1	Vertices	1 1 0 1 0 0	-20.71
		Middle	0 0 1 1 0 0	
	Ring 2	Vertices	0 0 0 1 0 1	
		Middle	0 1 1 1 0 1	
CHA	Ring 1	Vertices	1 1 1 1 1 0	-22.96
		Middle	1 0 1 1 0 1	
	Ring 2	Vertices	1 1 0 1 1 0	
		Middle	1 1 1 0 0 1	
CHA	Ring 1	Vertices	1 1 0 1 1 0	-20.24
		Middle	1 0 0 1 1 1	
	Ring 2	Vertices	1 0 0 1 1 1	
		Middle	1 1 1 0 0 1	

**Fig. 4** Normalized absolute power pattern of 12-element HA with 50 % thinning

Here, two cases are considered for two different values of inter-element spacing for the synthesis of HA. In the first case, spacing between elements is fixed at 0.5λ , and for the second case, spacing is optimized within the range $d_e \in [0.4, 0.7]$.

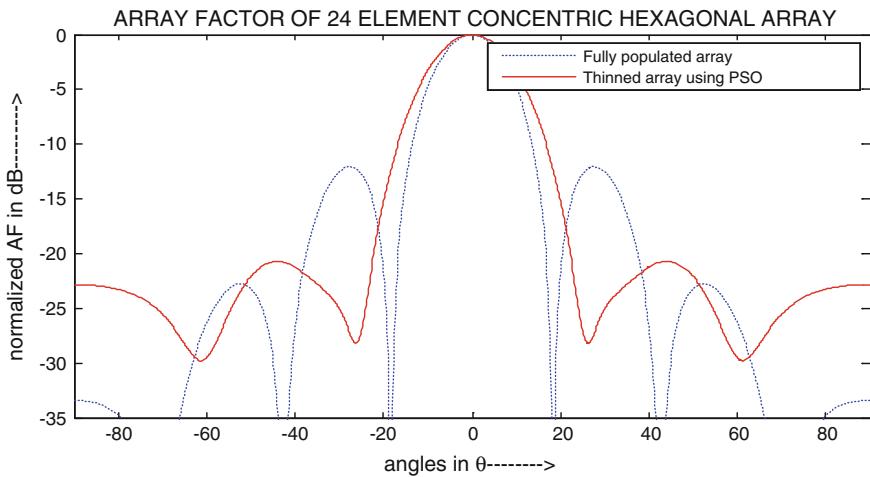


Fig. 5 Normalized absolute power pattern of 24-element CHA with 54.17 % thinning

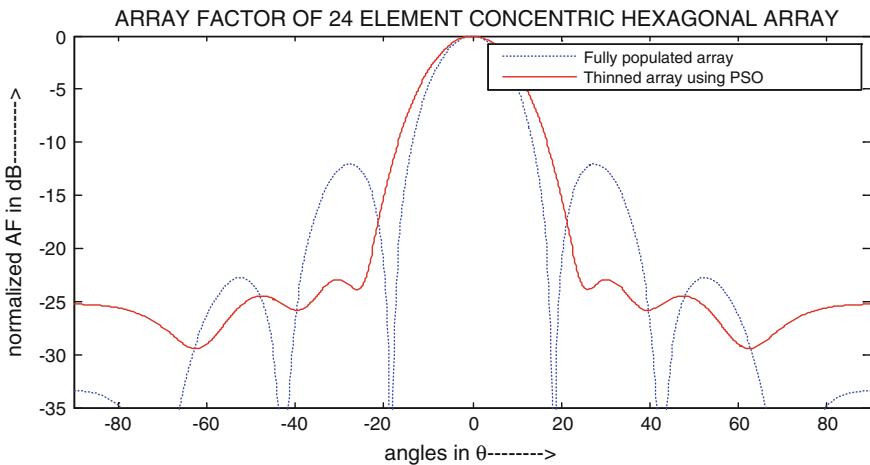


Fig. 6 Normalized absolute power pattern of 24-element CHA with 29.1 % thinning

Similarly, two cases are considered for two different values of inter-ring spacing for the synthesis of CHA. In the first case, spacing between two rings is fixed at 0.5λ , and for the second case, spacing is optimized within the range $d \in [0.4, 0.6]$. Here, inter-element spacing of the inner ring is kept fixed at 0.5λ .

The optimal non-uniformly excitation amplitudes for these cases of HA and CHA using PSO are shown in Tables 3 and 4, respectively.

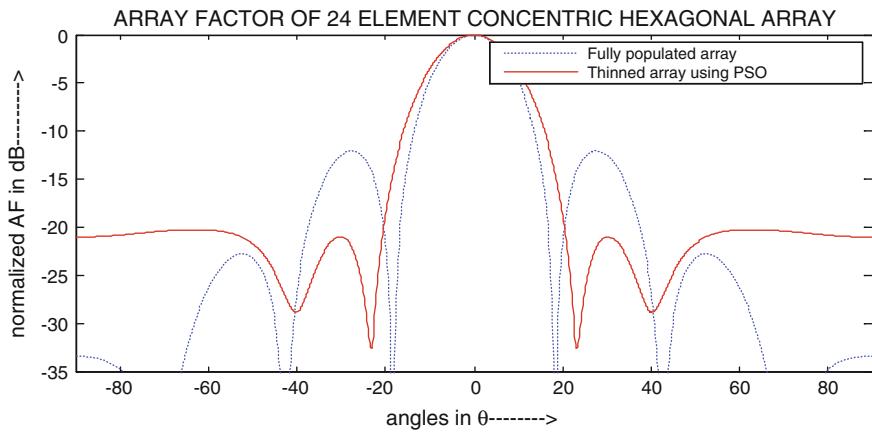


Fig. 7 Normalized absolute power pattern of 24-element CHA with 33.33 % thinning

Table 3 Excitation amplitude distribution using PSO for the two cases of 12-element HA

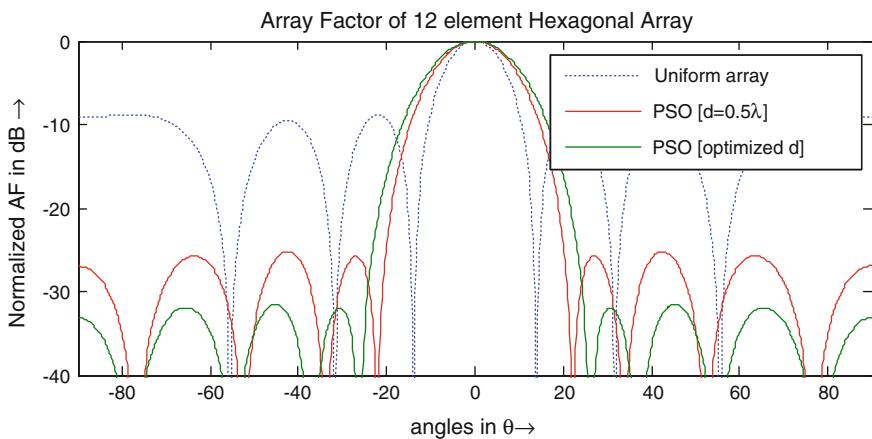
Inter-element spacing d_e (λ)	Placement of elements	Amplitude distributions			Max. SLL (dB)	FNBW in degree
Fixed $d_e = 0.5\lambda$	Vertices (A_n)	0.8642	0.1820	0.0123	-25.22	44.4
		0.9877	0.5328	0.0991		
	Middle (B_n)	0.9090	0.5110	0.5526		
		0.7569	0.0770	0.7078		
Optimized $d_e = 0.4754\lambda$	Vertices (A_n)	0.6954	0.4266	0.0893	-31.54	52.4
		0.9977	0.0003	0.4595		
	Middle (B_n)	1.0000	0.0000	0.2961		
		0.3763	0.0384	0.8952		

Figures 8 and 9 represents the normalized power patterns for different cases of the optimized HA and CHA, respectively. From Fig. 8, we can see that SLL is reduced to -25.22 dB by amplitude-only optimization and -31.54 dB with optimized inter-element spacing along with amplitude for a 12-element HA. From Fig. 9, also we can see that SLL is reduced to -24.68 dB by amplitude-only optimization and -27.17 dB with optimized inter-ring spacing along with amplitude for a double-ring 24-element CHA.

Figures 10 and 11 show the convergence curve using PSO for the synthesis of HA and CHA, respectively.

Table 4 Excitation amplitude distribution using PSO for the two cases of 24-element CHA

Inter-ring spacing d (λ)	Placement of elements		Amplitude distributions			Max. SLL (dB)	FNBW in degree
Fixed $d = 0.5\lambda$	Ring 1	Vertices (A_n)	0.8980	0.3925	0.4042	-24.68	41.6
			0.4347	0.9568	0.2037		
		Middle (B_n)	0.9995	0.3275	0.7274		
			0.5991	0	1.0000		
	Ring 2	Vertices (A_n)	0.8441	0.9098	0.0000		
			0.8987	0.5756	0.2475		
		Middle (B_n)	0.3351	0.3425	0.6908		
			0.5216	0.3708	0.4842		
Optimized $d = 0.5119\lambda$	Ring 1	Vertices (A_n)	0.7479	0.0436	0.4627	-27.17	44.0
			1.0000	0.4822	0.1678		
		Middle (B_n)	0.9898	0.9263	1.0000		
			0.9169	0.5823	1.0000		
	Ring 2	Vertices (A_n)	0.6172	0.6850	0.0022		
			0.5807	0.3568	0.2238		
		Middle (B_n)	0.1743	0.4009	0.4322		
			0	0.3833	0.6131		

**Fig. 8** Normalized absolute power patterns of the optimized HA

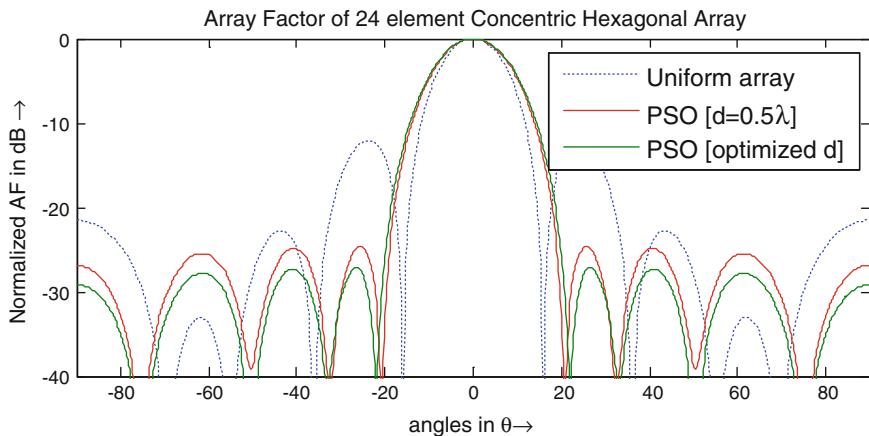


Fig. 9 Normalized absolute power patterns of the optimized CHA

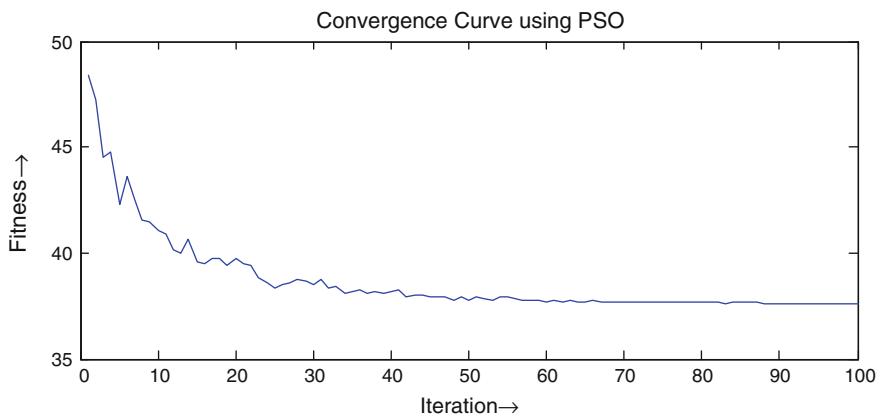


Fig. 10 Convergence profile for a 12-element HA using PSO

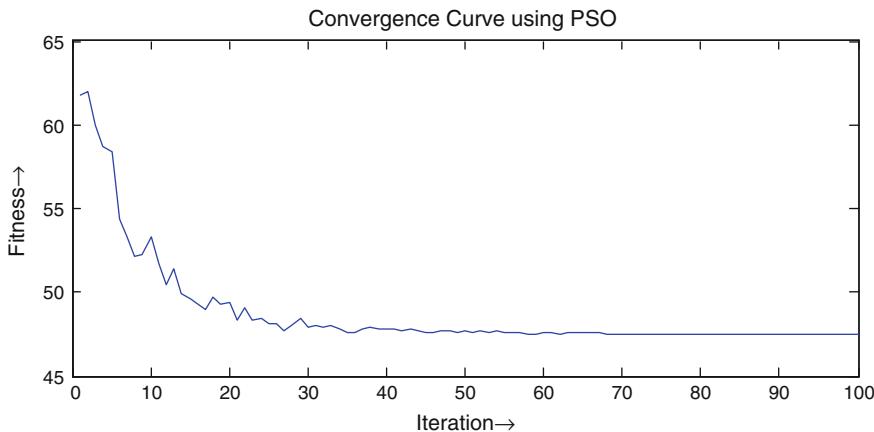


Fig. 11 Convergence profile for a 24-element CHA using PSO

5 Conclusions

In this paper, optimization of two hybrid antenna arrays called HA and CHA of uniformly excited isotropic antennas which can generate directive beam with minimum relative SLL is considered. Simulation results show that the number of effective antenna elements can be brought down from 12 to 6 totals with simultaneous reduction in SLL to less than -23 dB in HA. And in CHA, number of effective antenna elements can be brought down more than 50 % of totals with simultaneous reduction in SLL to less than -20 dB. Also, from results, we can see that SLL is reduced to less than -31 dB with optimized inter-element spacing along with amplitude for HA and less than -27 dB with optimized inter-ring spacing along with amplitude for a double-ring CHA using PSO. The performance of PSO shows the significant improvement for SLL with significant reduction in the number of elements which will reduce the cost of designing the antenna array substantially.

References

1. Balanis, C.A.: Antenna Theory and Design, 3rd edn. Wiley, New York (2005)
2. Chryssomallis, M.: Smart antennas. *IEEE Antennas Propag. Mag.* **42**(3), 129–136 (2000)
3. Ioannides, P., Balanis, C.A.: Uniform circular arrays for smart antennas. *IEEE Antennas Propag. Mag.* **47**(4), 192–206 (2005)
4. Lotfi, A.A., Ghiamy, M., Moghaddasi, M.N., Sadeghzadeh, R.A.: An investigation of hybrid elliptical antenna arrays. *IET Microw. Antennas Propag.* **2**(1), 28–34 (2008)
5. Mahmoud, K.R., El-Adway, M., Ibrahem, S.M.M., Basnel, R., Mahmoud, R., Zainud-Deen, S.H.: A comparison between circular and hexagonal array geometries for smart antenna systems using particle swarm algorithm. *Prog. Electromagn. Res. [PIER]* **72**, 75–90 (2007)

6. Dessouky, M., Sharshar, H., Albagory, Y.: Efficient sidelobe reduction technique for small-sized concentric circular arrays. *Prog. Electromagnet. Res. [PIER]* **65**, 187–200 (2006)
7. Bera, R., Roy, J.S.: Optimization of thinned elliptical antenna arrays using particle swarm optimization. *Proc. IEEE CODIS* 527–530, 28–29 Dec. 2012
8. Razavi, A., Forooraghi, K.: Thinned arrays using pattern search algorithms. *Prog. Electromagnet. Res. [PIER]* **78**, 61–71 (2008)
9. Haupt, R.L.: Thinned arrays using genetic algorithms. *IEEE Trans. Antennas Propag.* **42**(7), 993–999 (1994)
10. Schwartzman, L.: Element behavior in a thinned array. *IEEE Trans. Antennas Propag.* **15**(7), 571–572 (1967)
11. Mandal, D., Ghoshal, S.P., Bhattacharjee, A.K.: Design of concentric circular antenna array with central element feeding using particle swarm optimization with constriction factor and inertia weight approach and evolutionary programing technique. *J. Infrared Millimeter Terahz Waves* **31**(6), 667–680 (2010)
12. Eberhart, R.C., Shi, Y.: Particle swarm optimization: Developments, applications and resources. *Proc. Congr. Evol. Comput.* **1**, 81–86 (2001)
13. Jin, N., Rahmat-Samii, Y.: Advances in particle swarm optimization for antenna designs: Real-number, binary, single-objective and multi-objective implementations. *IEEE Trans. Antennas Propag.* **55**(3), 556–567 (2007)
14. Mangoud, M.A., Elragal, H.M.: Antenna array pattern synthesis and wide null control using enhanced particle swarm optimization. *Prog. Electromagnet. Res. B* **17**, 1–14 (2009)

Bearing-Only Tracking Using Sparse-Grid Gauss–Hermite Filter

Rahul Radhakrishnan, Shovan Bhaumik, Nutan Kumar Tomar
and Abhinoy Kumar Singh

Abstract In this paper, performance of sparse-grid Gauss–Hermite filter (SGHF) in bearings-only tracking (BOT) problem has been studied and compared with the performance of unscented Kalman filter (UKF), cubature Kalman filter (CKF), and Gauss–Hermite filter (GHF). The performance has been compared in terms of estimation accuracy and percentage of track loss, subjected to high initial uncertainty. It has been found that track loss of SGHF is less than all other quadrature filters with comparable estimation accuracy.

Keywords Bearing-only tracking · Nonlinear filtering · Gauss–Hermite filter · Sparse-grid Gauss–Hermite filter · Smolyak rule

1 Introduction

Bearings-only tracking is an important field of study in nonlinear filtering applications. This problem has huge importance in tracking and estimating the position and velocity of a moving target using passive sensors. This makes it an inevitable part in defense applications like tracking an aircraft, warships, submarines, and torpedos [1], as it does not reveal the identity of the tracker. The inherent nonlinear

R. Radhakrishnan (✉) · S. Bhaumik · N.K. Tomar · A.K. Singh

Indian Institute of Technology Patna, Patna, Bihar, India

e-mail: rahul.pee13@iitp.ac.in

URL: <http://www.iitp.ac.in>

S. Bhaumik

e-mail: shovan.bhaumik@iitp.ac.in

N.K. Tomar

e-mail: nktomar@iitp.ac.in

A.K. Singh

e-mail: abhinoy@iitp.ac.in

property of the measurements and the added criteria of observability problem make the bearings-only tracking (BOT) problem a complex one.

BOT problems have been widely discussed in the literature. Most of the research in this field has concentrated on the tracking of nonmaneuvering targets [2, 3]. In these works, tracking of a constant-velocity target is mentioned, where all the bearing measurements are obtained from a single moving observer. It has been found that the resulting dynamic model is not observable prior to an observer maneuver [4]. Tracking of maneuvering targets has also been explored in the literature [5, 6]. All these works modeled the target dynamics by multiple switching models using a jump Markov system framework, while the measurements are collected from a stationary or possible multiple moving observers.

Bayesian approach provides a general framework for solving the BOT problem. But due to the nonlinearities in the system measurement model, the optimal solution from Bayesian filtering framework becomes mathematically intractable. Extended Kalman filter (EKF) was the first introduced suboptimal algorithm, but it suffers from poor track accuracy and also track divergence due to linearization of measurement model [7]. Succeeding EKF, a variety of nonlinear filtering algorithms were introduced such as the unscented Kalman filter (UKF) [8], cubature Kalman filter (CKF) [9], and cubature quadrature Kalman filter (CQKF) [10]. In a quest to achieve more accuracy, another filter named Gauss–Hermite filter (GHF) was introduced [11]. GHF makes use of Gauss–Hermite quadrature rule for univariate systems, and these points are extended to multidimensional domain by using the product rule. This increases the number of multivariate points exponentially and hence suffers from the *curse of dimension* problem.

Sparse-grid Gauss–Hermite filter (SGHF) achieves similar accuracy levels as compared to GHF, with reduced computational load [12]. In this technique, the univariate quadrature rule is extended to multidimension with the help of the Smolyak rule. This theory remarkably reduces the computational load of the algorithm. Instead of using the moment matching method for defining the univariate quadrature points as described in [12], Golub’s technique can also be used [11]. In this case, the resulting filter was named as sparse-grid Gauss–Hermite filter (SGHF).

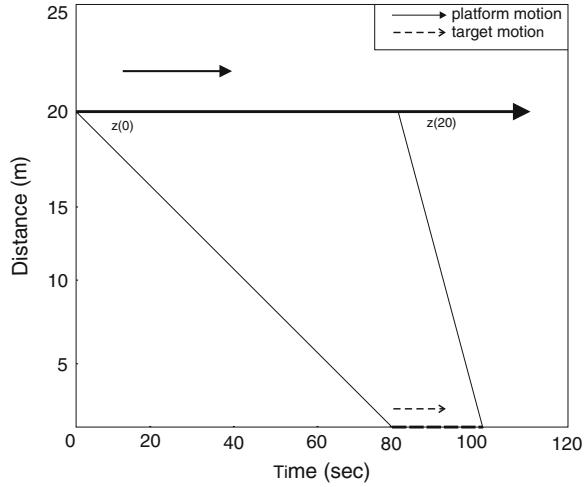
In this paper, performance of SGHF for a BOT problem is studied in terms of computational load, percentage of track loss, and accuracy. Results are compared with other quadrature filters like UKF, CKF, CQKF, and GHF.

2 Problem Formulation

In this problem, the target is assumed to be moving in a horizontal plane (Fig. 1) and platform, in the same vertical plane. Both are assumed to have a near-constant-velocity model.

Process model: The target dynamics can be expressed in discrete state space representation as [13, 14]:

Fig. 1 Target and tracking platform kinematics



$$X_{k+1} = F_k X_k + G_k w_k \quad (1)$$

where $X_k = [x_{1,k} \ x_{2,k}]^T$, $F_k = [1 \ T; \ 0 \ 1]$, and $G_k = [T^2/2 \ T]^T$. Here, $x_{1,k}$ is the position along the x -axis in m and $x_{2,k}$ is the velocity in m/s. $w_k \in \mathbb{N}(0, q)$, where \mathbb{N} is the normal distribution and T is the sampling time. The noise covariance matrix for target motion is given as

$$Q_k = G_k G_k^T = [T^2/2 \ T]^T [T^2/2 \ T] q = \begin{bmatrix} T^4/4 & T^3/2 \\ T^3/2 & T^2 \end{bmatrix} q. \quad (2)$$

The discrete time equations for the tracking platform motion may be represented as [13, 14]:

$$x_{p,k} = \bar{x}_{p,k} + \Delta x_{p,k} \quad k = 0, 1, \dots, n_{\text{step}} \quad (3)$$

$$y_{p,k} = \bar{y}_{p,k} + \Delta y_{p,k} \quad k = 0, 1, \dots, n_{\text{step}} \quad (4)$$

where $\bar{x}_{p,k}$ and $\bar{y}_{p,k}$ are the known average platform position coordinates and $\Delta x_{p,k} \in \mathbb{N}(0, r_x)$ and $\Delta y_{p,k} \in \mathbb{N}(0, r_y)$.

Measurement model: The measurement model can be represented as [15]:

$$z_{m,k} = \tan^{-1} \left(\frac{y_{p,k}}{x_{1,k} - x_{p,k}} \right) + v_{s,k} \quad (5)$$

where $v_{s,k} \in \mathbb{N}(0, r_s)$. The random component of platform motion induces additional measurement error, which is nonadditive and already embedded in Eq. (6).

The above effect can be approximated as additive noise by expanding the nonlinear measurement equations as

$$z_{m,k} = h[x_{p,k}, y_{p,k}, x_{1,k}] + v_{s,k} \approx h[\bar{x}_{p,k}, \bar{y}_{p,k}, x_{1,k}] + v_k \quad (6)$$

where v_k is the equivalent additive measurement noise (with variance R_k) given approximately by small perturbation theory as

$$v_k \approx \frac{\bar{y}_p \Delta x_p + \{x_{1,k} - \bar{x}_{p,k}\} \Delta y_p}{[x_{1,k} - \bar{x}_{p,k}]^2 + \bar{y}_{p,k}^2} + v_{s,k}. \quad (7)$$

R_k is calculated considering $\Delta x_{p,k}$, $\Delta y_{p,k}$, and v_k to be mutually independent, given by

$$R_k = E[v_k^2] = \frac{\bar{y}_{p,k}^2 r_x + [x_{1,k} - \bar{x}_{p,k}]^2 r_y}{\left\{ [x_{1,k} - \bar{x}_{p,k}]^2 + \bar{y}_{p,k}^2 \right\}^2} + r_s. \quad (8)$$

3 Sparse-Grid Gauss–Hermite Filter

In SGHF, the single-dimensional quadrature rule is extended to multidimension by using the Smolyak rule. It can reduce the computational load considerably when compared to the product rule used in GHF.

Smolyak rule: Any integral of the form [12]

$$I_{n,L}(f) = \int_{\mathbb{R}^n} f(x) \mathbb{N}(x; 0, I_n) dx$$

can be approximated numerically as

$$I_{n,L}(f) \approx \sum_{q=L-n}^{L-1} (-1)^{L-1-q} C_{L-1-q}^{n-1} \sum_{\Xi \in N_q^n} (I_{l_1} \otimes I_{l_2} \otimes \cdots \otimes I_{l_n})(f), \quad (9)$$

where $I_{n,L}$ represents the numerical evaluation of n -dimensional integral with the accuracy level L which means that the approximation is exact for all the polynomials having degree up to $(2L - 1)$. C stands for the binomial coefficient, i.e., $C_k^n = n!/k!(n-k)!$, I_{l_j} is the single-dimensional quadrature rule with accuracy level $l_j \in \Xi$, i.e., $\Xi \triangleq (l_1, l_2, \dots, l_n)$, \otimes stands for the tensor product, and N_q^n is a set of possible values of l_j given as

$$N_q^n = \begin{cases} \Xi : \sum_{j=1}^n l_j = n + q \end{cases} & \text{for } q \geq 0 \\ = \emptyset & \text{for } q < 0. \end{cases} \quad (10)$$

Equation (9) can be written as

$$\begin{aligned} I_{n,L}(f) \approx & \sum_{q=L-n}^{L-1} (-1)^{L-1-q} C_{L-1-q}^{n-1} \sum_{\Xi \in \mathbf{N}_q^n} \sum_{q_{s_1} \in X_{l_1}} \sum_{q_{s_2} \in X_{l_2}} \\ & \dots \sum_{q_{s_n} \in X_{l_n}} f(q_{s_1}, q_{s_2}, \dots, q_{s_n}) w_{s_1} w_{s_2} \dots w_{s_n}, \end{aligned} \quad (11)$$

where X_{l_j} is the set of quadrature points for the single-dimensional quadrature rule I_{l_j} , $[q_{s_1}, q_{s_2}, \dots, q_{s_n}]^T$ is a sparse-grid quadrature point, i.e., $q_{s_j} \in X_{l_j}$ and w_{s_j} is the weight associated with q_{s_j} . Some quadrature points may occur multiple times, which could be counted only once and by adding their weights. The final set of the SGQ points can be represented as

$$X_{n,L} = \bigcup_{q=L-n}^{L-1} \bigcup_{\Xi \in \mathbf{N}_q^n} (X_{l_1} \otimes X_{l_2} \otimes \dots \otimes X_{l_n}), \quad (12)$$

where \cup represents the union of the individual SGQ points.

4 Simulations

The filtering characteristics of SGHF is studied in detail and compared with various other suboptimal filters. The (unknown) true initial condition and the known noise variance are assumed to be $X_0 = [80 \ 1]^T$ and $q = 0.01 \text{ m}^2/\text{s}^4$, respectively. The mean positions of the platform are $\bar{x}_{p,k} = 4 \text{ kT}$ and $\bar{y}_{p,k} = 20$, and its noise variances are selected as $r_x = 1 \text{ m}^2$ and $r_y = 1 \text{ m}^2$, respectively. For the measurement model, the variance associated with the noisy measurements is selected as $(r_s = (3^\circ)^2)$.

Usually, filter initialization is done from the first two measurements. The latest bearing measurement defines the initial position estimate, and the difference of two bearing measurements provides the estimation of initial velocity. The initial position estimate thus obtained may be shown to have a covariance of

$$P_{11,0} = r_x + \frac{r_y}{\tan^2 z} + \frac{\bar{y}_p^2}{\sin^4 z} r_s \quad (13)$$

Here, we consider the initial position and velocity estimates as $x_{1,0} = \mathbb{N}(x_{1,0}, P_{11,0})$ and $x_{2,0} = \mathbb{N}(x_{2,0}, P_{22,0})$, where $P_{22,0} = 1$. The off-diagonal terms $P_{12,0}$ and $P_{21,0}$ are taken as zero.

Results were compared with a number of 1,000 Monte Carlo runs for 20 time steps, and the value of each time step is $T = 1$ s. The scaling parameter κ in UKF algorithm is taken according to the condition $n + \kappa = 3$. The accuracy levels of GHF are selected as $m = 5$ and that of SGHF, $L = 3$. Root mean square error (RMSE) plots of position and velocity estimates are shown in Figs. 2 and 3. It can be inferred from the plots that the SGHF shows comparable accuracy levels with respect to GHF and all other filters. The track loss criteria are defined as $|x_{20} - \hat{x}_{20}| < X_{\text{end}}$, where $X_{\text{end}} = 15$ m and the subscript 20 denotes the 20th time step. The percentage of track loss was calculated for 10,000 runs and has been displayed in Table 1. From Table 1, it can be observed that the track loss count of SGHF is less than all other filters with the UKF possessing highest number of track loss cases. The main advantage of SGHF over GHF is that it consumes less computational time. This is

Fig. 2 RMSE of position for 1,000 MC runs

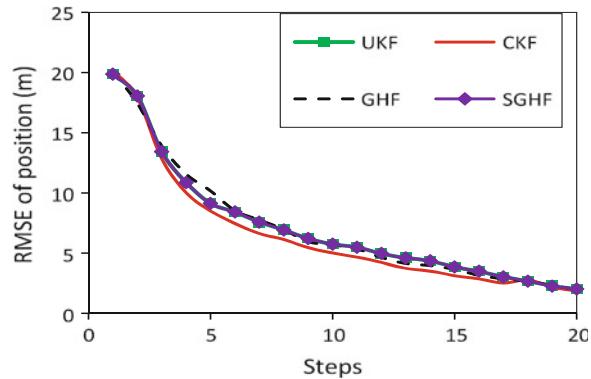


Fig. 3 RMSE of velocity for 1,000 MC runs

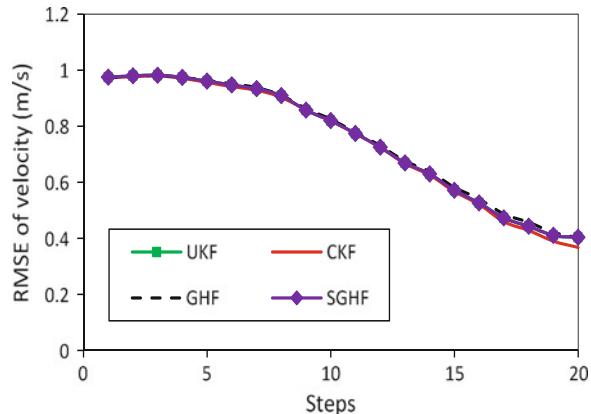


Table 1 Track loss and relative computational time for various filters

Filters	% Track loss	Relative computational time
UKF	3.07	1
CKF	1.98	0.89
GHF	1.76	3.34
SGHF	1.73	1.27

evident from Table 1, where SGHF consumes much less than half of the time consumed by GHF. Thus, we conclude that SGHF can be considered as a highly efficient filtering algorithm for BOT problem.

5 Conclusion

In this work, performance of sparse-grid Gauss–Hermite filter for a particular BOT problem is studied. The performance of the filter was studied and compared with other filters by using RMSE criteria. RMSE plots illustrate that SGHF shows comparable accuracy levels with GHF, with much less computational time. The number of track losses incurred in SGHF is also comparable to that of GHF. Since it possess comparable accuracy levels with respect to CKF and UKF and consumes much less computational time with respect to GHF, it can be highly preferred for onboard applications.

References

1. Karlsson, R., Gustafsson, F.: Recursive Bayesian estimation: bearings-only applications. In: IEE Proceedings-Radar, Sonar and Navigation, vol. 152, pp. 305–313. IET, Stevenage (2005)
2. Aidala, V.J.: Kalman filter behavior in bearings-only tracking applications. IEEE Trans. Aerosp. Electron. Syst. **AES-15**(1), 29–39 (1979)
3. Farina, A.: Target tracking with bearings-only measurements. Sig. Process. **78**(1), 61–78 (1999)
4. Nardone, S.C., Aidala, V.J.: Observability criteria for bearings-only target motion analysis. IEEE Trans. Aerosp. Electron. Syst. **AES-17**(2), 162–166 (1981)
5. Ristic, B., Arulampalam, M.S.: Tracking a manoeuvring target using angle-only measurements: algorithms and performance. Sig. Process. **83**(6), 1223–1238 (2003)
6. Arulampalam, M.S., Ristic, B., Gordon, N., Mansell, T.: Bearings-only tracking of manoeuvring targets using particle filters. EURASIP J. Appl. Sig. Process **2004**, 2351–2365 (2004)
7. Julier, S.J., Uhlmann, J.K.: A new extension of the Kalman filter to nonlinear systems. In: Proceedings of SPIE: Signal Processing, Sensor Fusion, and Target Recognition VI, p. 182 (1997)
8. Julier, S.J., Uhlmann, J.K.: Unscented filtering and nonlinear estimation. Proc. IEEE **92**(3), 401–422 (2004)

9. Arasaratnam, I., Haykin, S.: Cubature Kalman filters. *IEEE Trans. Autom. Control* **54**(6), 1254–1269 (2009)
10. Bhaumik, S., Swati: Cubature quadrature Kalman filter. *Sig. Process. IET* **7**(7), 533–541 (2013)
11. Arasaratnam, I., Haykin, S., Elliott, R.J.: Discrete-time nonlinear filtering algorithms using Gauss-Hermite quadrature. *Proc. IEEE* **95**(5), 953–977 (2007)
12. Jia, B., Xin, M., Cheng, Y.: Sparse-grid quadrature nonlinear filtering. *Automatica* **48**(2), 327–341 (2012)
13. Lin, X., Kirubarajan, T., Bar-Shalom, Y., Muskell, S.: Comparison of EKF pseudo measurement and particle filter for a bearing only target tracking problem. In: Proceedings of SPIE: Signal and Data Processing of Small Targets (2002)
14. Chalasani, G., Bhaumik, S.: Bearing only tracking using Gauss-Hermite filter. In: 7th IEEE Conference on Industrial Electronics and Applications (ICIEA), pp. 1549–1554. IEEE, New York (2012)
15. Bar-Shalom, Y., Li, X.R., Kirubarajan, T.: Estimation with applications to tracking and navigation: theory algorithms and software. Wiley, London (2004)

Novel Design and Implementation of Passive Infrared Sensor in Steel Industry Automation

Basava Naga Girish Koneru, Vijaya Krishna Teja Mantripragada, Prithvi Raj Pani and G. Lakshminarayanan

Abstract In the present-day scenario, industries are being completely automated in order to achieve better products at a faster rate. Automation also includes efficient monitoring system to enhance the safety standards. This paper presents how passive infrared (PIR) sensor, when implemented in combination with GSM technology, would smoothen the manufacturing process. Also, as these industries deal with very sophisticated machineries on a large scale, there should be no compromise, when it comes to the safety of working personnel. We have also worked out how PIR sensor along with a suitable optical filter can be implemented to achieve the desired safety levels in a steel industry.

Keywords Industrial automation · Industrial safety · Passive infrared (PIR) sensor · GSM SIM900 module · Optical filter · 8051 microcontroller · Foolproof system

1 Introduction

With rapid advancement in the field of automation, practically every sphere of life is getting automated. Moreover, with increased competition in the consumer market, no industry can afford to fall short of their productivity target. Keeping this in mind,

B.N.G. Koneru (✉) · V.K.T. Mantripragada · P.R. Pani · G. Lakshminarayanan
Electronics and Communication Engineering Department, National Institute
of Technology, Trichy, Tamil Nadu, India
e-mail: 108112040@nitt.edu

V.K.T. Mantripragada
e-mail: 108112099@nitt.edu

P.R. Pani
e-mail: paniprithviraj@gmail.com

G. Lakshminarayanan
e-mail: laksh@nitt.edu

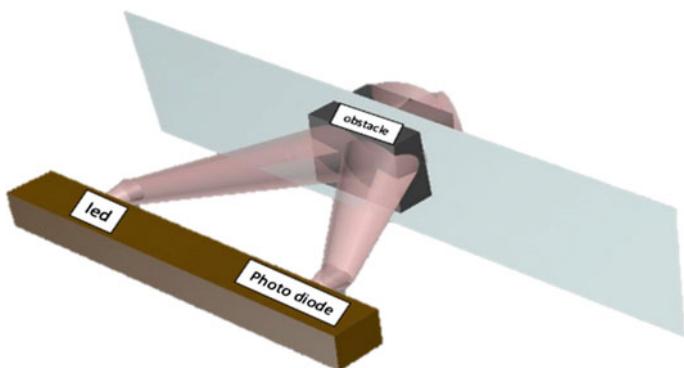


Fig. 1 Conventional IR sensor technology

in this paper, we have worked on developing the passive infrared (PIR) sensor module to enhance the safety and increase efficiency in the manufacturing process. This idea of using PIR sensor in industrial automation has been implemented with wire rod mill section in a steel industry. We have used the PIR sensor technology to improve upon the existing conventional infrared (IR) sensor technology for monitoring the input and output processes in the steel industry. The existing technology involves the usage of active IR transmitter-receiver pair, which makes the use of reflection phenomena of IR waves to monitor the movement of charge (Fig. 1).

The proposed model of PIR sensor outshines the existing mechanism by being more sensitive to the movement of objects (input feed or finished product or any undesirable intervention). In this work, two PIR sensors are used simultaneously at two different locations. Also, we made use of GSM SIM900 module, which faithfully communicates the on-field scenario, hence ensuring a foolproof automated monitoring of the conveyor system (charge or finished product). Besides using the GSM SIM900 module, we also did make use of optical filters (band pass and band rejection), which ensures a safe process guarding against foreign materials. Hence, our module that can be termed as ‘smart industrial application’ enhances efficiency of the process and safety standards simultaneously (Fig. 2).

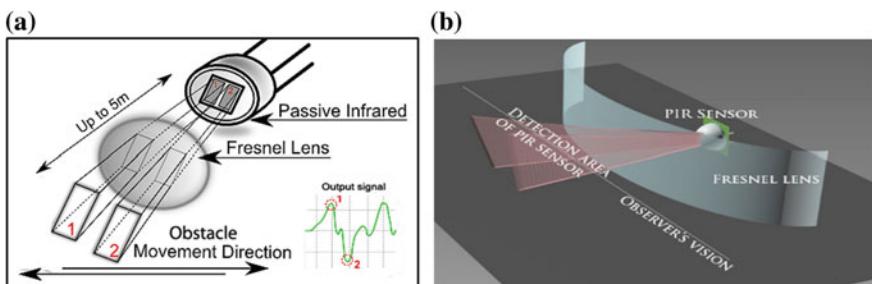


Fig. 2 a, b Working of PIR sensor

2 Existing System

Previous works in this area of PIR sensors have been limited to domestic surveillance systems. They have used the same basic functionality of PIR sensor of sensing the change in the infrared rays to detect any human intrusion and other phenomena [1–6]. These works were successfully applied in domestic households, ATMs, banks, and shopping malls. But in this paper, we took a step further, exploiting the PIR sensor technology to industrial scenario. In our work, we have used PIR sensor as both monitoring unit and productivity enhancer simultaneously. Brief descriptions of some of the works are given below:

In [7], an ARM 7-based automation system has been suggested in conjunction with the wireless communication technology using ZigBee and RS-485 protocol for serial communication. The proposed system makes use of multiple sensors and wireless communication technology in order to control a LED light according to user's state and surroundings. Basically, the authors' proposal is related to domestic applications.

The work [8] has presented an elegant application to PIR sensor by implementing it beyond domestic applications. The paper has presented the PIR sensor detection to be specific to the human beings, and moreover, the DTMF technology has been made use of. But the drastic development of technology has probed us to efficiently utilize the GSM technology for a better performance.

In the work [9], a single PIR sensor was used to detect the human infrared signal. Here, the wavelet packet entropy (WPE) was used for extracting the features of the infrared signals received from the PIR sensor. In addition to this, least square support vector machine (LS-VSM) classifier was used, as a classifier. The basic principle of this paper is same as the proposed one, but its application is restricted to human detection alone. Also, it has been implemented in a way different from the classical one.

3 Proposed System—Hardware Implementation

In our proposed work, the hardware module comprises P89V51RD2BN (8051) microcontroller, two PIR sensors, MAX232 IC, DB9 cables, and GSM SIM900 modules. All of these modules are connected as per the standard circuit diagram, on a general-purpose board. Power is given by a 9-V battery (in the real-time industrial scenario, it can be replaced by ac-to-dc adaptor) and is regulated by using IC7805 voltage regulator. An LED across a $330\text{-}\Omega$ resistor is connected at the input terminal of the voltage regulator, to indicate the power supply (Fig. 3).

When the circuit is powered on, both the PIR sensors take 10–60 s to calibrate themselves with the existing environmental radiations. After getting calibrated, when there is a change of infrared radiation (induced due to movement of feed at input side or finished product at output side) in the field of view of the first PIR

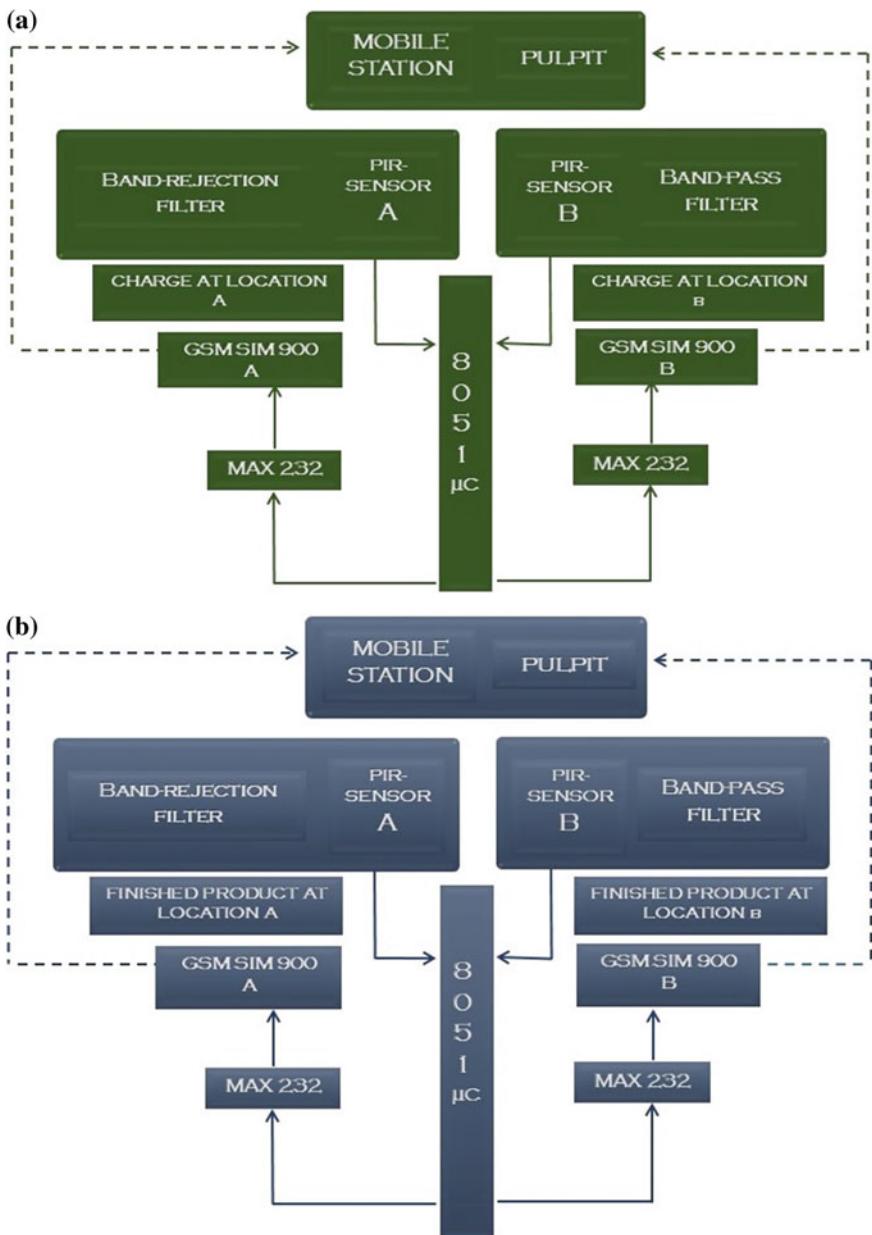


Fig. 3 **a** Proposed block diagram at the input section. **b** Proposed block diagram at the output section

sensor (A), it acts as the input for the band-pass optical filter placed in front of it. This band-pass optical filter chooses the preset wavelength (i.e., wavelength of the feed or the finished product) from the incoming radiation and sends it to the PIR

sensor (A). This sensor, upon detecting the change caused in the environmental radiation due to the movement of the charge (or the finished product), triggers the circuit connected across its output. The immediate neighbor of the PIR sensors is the microcontroller P89V51RD2BN. Upon the triggering of the PIR sensor (A), one of the pins of the microcontroller gets high and sends the signal to the GSM SIM900 (A) connected via MAX232 IC, stating the presence of the desired elements upon the conveyor belt.

Further as the conveyor belt moves forward, it causes a similar change of infrared radiations in the field of view of second PIR sensor (B). Here, a band-reject filter is placed in front of the PIR sensor (B), which rejects the band, characteristic of the charge (or the finished product). Therefore, this sensor does not trigger the output as long as there is the required material on the conveyor belt. But if there is anything other than that, this PIR sensor (B) triggers the microcontroller connected across its output. Then, the microcontroller triggers another GSM SIM900 (B) module which sends the message that there is something other than the desired charge (or the finished product) upon the conveyor belt and also this triggers an alarm in the pulpit. This alerts the staff in the control cabin who would then halt the process and issue necessary action.

Summarizing, the entire module is implemented in two stages. The first stage (PIR-GSM pair A) supports the efficiency of the process, by reducing the unwanted delays. Also, the first stage along with the second stage enhances the safety levels of working personnel as well as the industrial machineries, by halting the process in case of any unwanted interventions (Fig. 4).

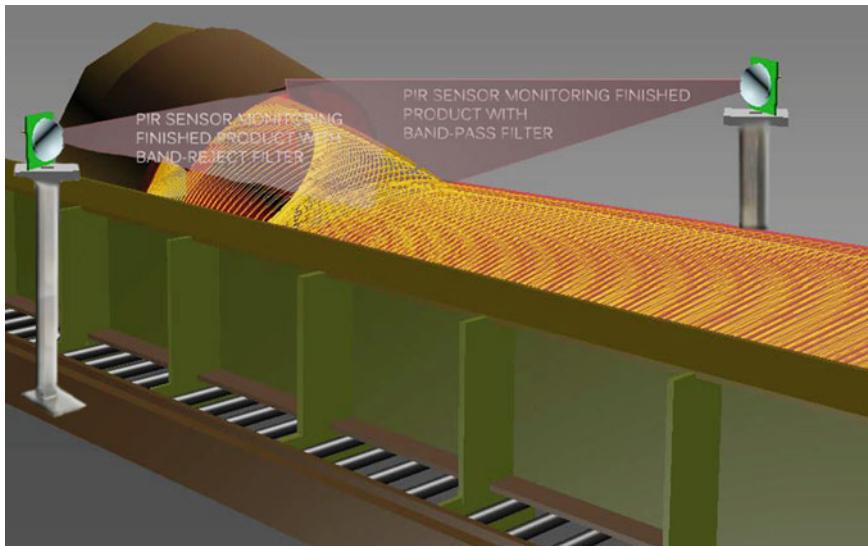


Fig. 4 Model of the proposed system

4 Circuit Description

The entire module consists of a general-purpose board which includes PIR sensors and GSM SIM900 module, MAX232 IC, and DB9 cables. Brief description about each of the elements of the circuit is as follows (Fig. 5):

4.1 Microcontroller

Microcontroller is used for creating a communication channel between the PIR sensor and the GSM module. P89V51RD2BN microcontroller, manufactured by Philips, was used here. This is a 40-pin IC with 64 KB on chip flash memory, 1,024 bytes of RAM data, four input–output ports. The operating voltage is 5 V at a clock frequency of 11.0592 MHz.

4.2 GSM Module

A global system for mobile communication (GSM) network consists of multiple components (such as transceivers, controllers, switches, routers, and registers), facilitating transmission or reception of signaling and traffic messages. It is being used here for communicating the status of feeding process to the control room. Subscriber identity module (SIM900) variant is being used here.

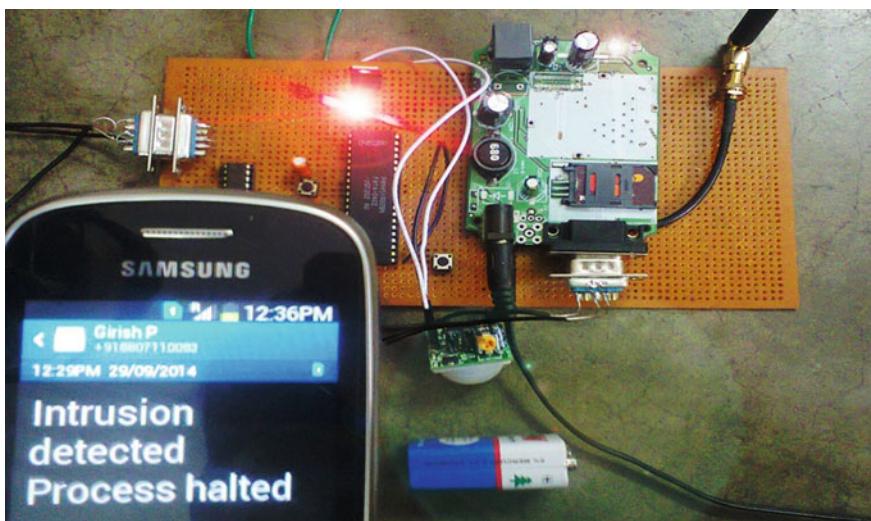


Fig. 5 Snapshot of working model

4.3 PIR Sensor

PIR sensors are fundamentally made up of a pyro electric sensor. It can detect various levels of infrared radiation. PIR sensors are mainly used in applications, involving detection of change in surrounding environment.

PIR sensors generally have 3 pins corresponding to ground, power supply, and output, respectively. Since the PIR functions as a digital output, we need to check for the pin to flip high or low on a single I/O pin. Once the PIR gets calibrated to the static environment (which takes a couple of seconds), the output remains low, until a motion is detected. The settling time is attributed to the fact that the PIR sensor needs some time to study the pre-existing domain of the nature. When the sensor detects an obstacle (feed/finished product or any intervention), it gives the digital output in the form of logic 1. Now, the message is sent to the control room via GSM module. The PIR sensor can monitor the flow of feed up to the distance of around 6 m with 120° field of view (Fig. 6).

4.4 Optical Filter

According to laws of thermal radiation, every object above 0 °C emits radiations of various wavelengths characteristic of each body. As the temperature of the material varies, the emitted wavelength of radiation also varies. These optical filters selectively allow a particular range of frequencies to the first PIR sensor (A). This helps us to detect the presence of the charge or the finished product alone. Then, as the charge or the finished product proceeds, it crosses the field of second PIR sensor (B). An optical band-reject filter precedes this PIR sensor. This setup helps us to detect the presence of any object other than the charge or finished product.

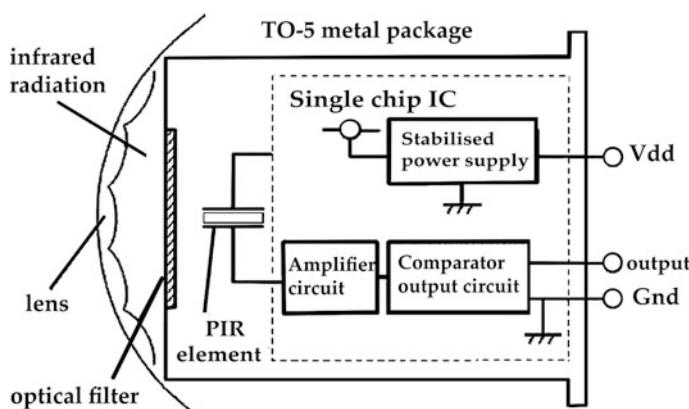


Fig. 6 PIR sensor internal circuitry

5 Future Enhancements

The use of PIR sensor can be also replaced by using IR Web cameras which take continuous snaps of the process. At the control panel, necessary action can be taken if any discrepancies are found in the images obtained from the cameras.

Here, GSM SIM900 has been used, which entails us the provision of Ethernet connection, thus making it adaptive for live monitoring. Industrial safety can also be enhanced further by immediately sending information based on the PIR module if there is any human intervention by using a wireless computer network for efficient solutions. Also, we have only taken one specific section (Wire Rod Mill) of a steel industry for illustration of our idea. But this can also be implemented in a different kind of industries (say automobile, chemical, food and agriculture, etc.) depending on the requirement.

6 Conclusion

This paper presents a novel method of enhancing the safety standards, also adding to the efficiency of the manufacturing process. A detailed account of the working and implementation of the PIR sensor in the steel industry has been presented. The complete description of the implementation of PIR sensor proves it to be a worthy candidate in the present industrial scenario. Moreover, this stands out to be unique owing to its three-dimensional field of view. Also, this module consumes low power compared to the conventional IR sensors because once the PIR sensor is calibrated, it does not consume further power to sustain itself. Also, this sensor helps in developing a foolproof system because of its insensitivity to the common dust which is unavoidable in the industrial environment.

Acknowledgments The authors deeply thank the Department of Electronics and Communication Engineering, National Institute of Technology, Tiruchirappalli, for providing us with the facilities for working on this project. Also, we are grateful to industrial experts and technocrats, who have been a source of encouragement and inspiration to make this attempt in the improving industrial automation.

References

1. Wang, L., Gong, W., He, L., Xiao, H., Huang, Y.: Human motion recognition using pyroelectric infrared signal. *J. Optoelectron. Laser* **21**(3), 440–443 (2010)
2. Fang, J.S., Hao, Q., Brady, D.J., Guenther, B.D., Hsu, K.Y.: Real-time human identification using a pyro-electric infrared detector array and hidden Markov models. *Opt. Express* **14**(15), 6643–6658 (2006)
3. Turaga, P., Chellappa, R., Subrahmanian, V.S., Udrea, O.: Machine recognition of human activities: A survey. *IEEE Trans. Circuits Syst. Video Technol.* **18**(11), 1473–1488 (2008)

4. Hernandez, S., Frean, M.: Bayesian multiple person tracking using probability hypothesis density smoothing. *Int. J. Smart Sens. Intell. Syst.* **4**(2), 285–312 (2011)
5. Ji, X., Liu, H.: Advances in view-invariant human motion analysis: A review. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **40**(1), 13–24 (2010)
6. Sen, W., Ying-Wen, C., Ming, X.: Research of passive infrared wireless sensor network target tracking. *J. Transduct. Technol.* **21**(11), 1929–1934 (2008)
7. Sulthana, M., Rao, N.U.: An ARM-7 based embedded energy efficient LED lighting system for domestic applications with user satisfaction, *IEEE* (2014)
8. Venkatesan, V.S.: GSM controlled robotics. In: *Fourth International Conference on Advanced Computing and Communication Technologies*, IEEE (2014)
9. Wang, L.: Human infrared signal recognition using single PIR detector. In: *Fourth International Congress on Image and Signal Processing* (2011)
10. Hong, S.G., Kim, N.S., Kim, W.W.: Reduction of false alarm signals for PIR sensor in realistic outdoor surveillance. *ETRI J.* **35**(1), 80–88 (2013)
11. Abubeker, K.M., Edathala, J.J., Sebastian, S.: Some investigations on ‘PIR’ sensor and their application on significant energy savings in ATM counter. In: *Proceedings of 2013 IEEE Conference on Information and Communication Technologies (ICT 2013)*
12. Xie, D., Liu, H., Li, B., Zhou, Q., Yuan, X.: Target Classification Using Pyro-electric Infrared Sensors in Unattended Wild Ground Environment. *Int. J. Smart Sens. Intell. Syst.* (2013). ISSN 1178-5608
13. Mazidi, J.G., McKinlay, R.D., Mazidi, M.A.: The 8051 Microcontroller and System

A Comparative Study of Feature Selection and Machine Learning Methods for Sentiment Classification on Movie Data Set

C. Selvi, Chakshu Ahuja and E. Sivasankar

Abstract Sentiment analysis has become a leading research domain with the advent of Web 2.0 where Web users express their opinions in user forums, blogs, discussion boards, and review sites. The online information is considered to be a valuable source for decision making, improving the quality of service, and helping the service providers to enhance their competitiveness. Since the processing of high-dimensional text data is not scalable, different feature selection mechanisms are being used to confine the study to only most informative features. These features are then used to train the classifier to improve the accuracy of sentiment-based classification. This paper explores six feature selection mechanisms (IG, GR, CHI, OneR, Relief-F, and SAE) with five different machine learning classifiers (SVM, NB, DT, K-NN, and ME) thereby providing Accuracy, on the movie review data set for each. Comparative results show that Naive Bayes (NB) outperforms other classifiers and works better for Gain Ratio (GR) and Significance Attribute Evaluation (SAE) feature selection method.

Keywords Sentiment analysis · Sentiment classification · Feature selection · Machine learning · Movie domain

C. Selvi (✉) · C. Ahuja · E. Sivasankar

Department of Computer Science and Engineering, National Institute of Technology,
Tiruchirappalli 620015, Tamil Nadu, India
e-mail: 406113001@nitt.edu

C. Ahuja
e-mail: 106111026@nitt.edu

E. Sivasankar
e-mail: sivasankar@nitt.edu

1 Introduction

Opinion Mining (OM) or Sentiment Analysis (SA) is the recent technology, which processes vast amount of user-generated unformatted text content [1], that is the user's opinion, attitude, view, sentiment toward particular product, or event. The vast Web information is highly unstructured and scattered. The purpose of SA is to determine the overall opinion of the user review in terms of positive and negative. Two approaches to perform sentiment classification [2, 3] on text reviews include Machine Learning (ML) [1–15] and Lexicon-based approach or Semantic-Oriented (SO) [8] approach. This paper focuses on ML approach on lexical features, where the classifier models such as a Support Vector Machine (SVM), Naive Bayes (NB), Decision Tree (DT), K-Nearest Neighbor (K-NN), and Maximum Entropy (ME) have been trained using a collection of labeled text data. Then, the trained model is employed to predict the new text data as either positive or negative. Every sentiment word in the text is not exposed explicitly. Instead, some of the words are obscured in the subjective text. Therefore, sentiment classification needs in-depth analysis to find out the informative features in the text. This paper focuses on the importance of feature selection for sentiment classification and also shows the experimental comparison of several feature selection methods applied to ML. In the rest of this paper, Sect. 2 presents related work on sentiment analysis. Section 3 presents related feature selection, feature weighting, and classifiers used. Section 4 describes the analysis on the empirical results of work done. Lastly, Sect. 5 concludes this paper.

2 Related Work

Sentiment analysis has been considered to be an important research subject in many domains, including social media and product sites. Table 1 describes related research work on SA using different ML techniques for movie review data set [14]. The table also shows the various features considered for classification, the features that perform best for classification, and their accuracy. Extracting informative feature in appropriate number is an open issue in sentiment classification and much of the work has been done on simple and word N-gram features. Feature selection methods used for classification includes CHI squared, Information Gain (IG), Gain Ratio (GR), Relief-F, Document Frequency (DF), Mutual Information (MI) [6, 14], significance attribute evaluation [16], and Fisher discriminant ratio [5].

From the existing literature, it is clear that there is a scope for improving the classifier accuracy by reducing the number of features. This paper focuses the work done on accuracy improvement of ML concepts using different feature selection techniques on unigram of words.

Table 1 Related works for sentiment analysis using ML techniques

	Classifier model	Data source	Features considered	Best features	Accuracy
[3]	NB, ME, SVM	Movie reviews (MR) 700 (+ and -)	N-grams, POS (part-of-speech), position	N-grams	NB of unigram and POS TP: 81.5 % ME of unigram and bigram TP: 80.8 %, SVM of unigram TP: 82.9 %
[10]	NB, SVM	MR 1,000 (+ and -)	N-grams, position	N-grams	NB: 86.4 % SVM: 87.15 %
[4]	SVM, term weighting	MR 1,000 (+ and -)	GI, CTRW, adjective, valance shifters	Unigram with bigram, valance shifter	SVM: 85.9 %
[2]	NB, multinomial ME, SVM	MR 1,000 (+ and -)	Unigram, unigram and subjectivity analysis, bigram, adjectives	Unigram and subjectivity analysis	SVM and term weighting: 86.2 %
[7]	DT, SVM, NB	MR 1,000 (+ and -)	Unigram TP, unigram TF, unigram aggregate	Unigram TP and aggregate	ME of feature presence for MR: 87.40 %
[12]	SVM, cluster based method	Movie reviews, product review with 1,000 (+ and -)	Eigenvector obtained by eigen decomposition	Unigram TP and aggregate	ME for reviews from blog corpus: 90.25 %
[13]	SVM	Multi-domain data set (MDSD)–8,000, MR 1,000 (+ and -)	Unigram TP, TF, TF-IDF	Unigram TF-IDF	SVM: 77.0 % DT: 69.3 % NB: 77.5 %
[8]	NB, ME, SVM, meta classifier	Movie reviews, product review with 1,000 (+ and -)	POS1-Adj + Adv, POS2-Verb, POSS3-Noun, joint POS, unigram, bigram, word dependencies	POS1, joint POS, joint WR features	Amazon product review: 93.7 % MDSD: 96.40 % Movie review: 96.90 %
[9]	SVM, ANN	Movie, GPS, Camera, Books	Unigram TF-IDF	Unigram TF-IDF	ANN for movie: 86.5 % Camera: 81.8 % GPS: 87.3 %, Book: 90.3 %

3 Methodology

The algorithm for classifying text review:

- Step 1: Collect labeled text reviews
- Step 2: Apply tokenization on the text review to get unigram
- Step 3: Apply feature selection ranking method on unigram feature and find the weight of each feature and rank them in descending order
- Step 4: Select the top k features, where $1,000 \leq k \leq 5,000$
- Step 5: Build a k -fold classifier model with selected features using feature presence weighting scheme
 - Function of k -fold classifier model
 - Partition the data set into k equal size subsample set
 - Train a classifier model with $k-1$ subsample set and validate a model with one subsample set
 - Repeat the above step for k times
 - The performance of the classifier model on entire data set is calculated by averaging the above-obtained result
- Step 6: Repeat Step 5 for different ranges of selected features and analyze the accuracy of classifier
- Step 7: Find the classifier with highest accuracy for the best range of features selected by the feature selection method

All the reviews are labeled with two classes (+ve and -ve). Let f_i ($1 \leq i \leq m$) be the vector of m unigram features for all reviews used for review classification. This section describes various feature filtering mechanisms, feature weighting schemes, and classifier techniques used to categorize the review as a positive and negative review.

3.1 Feature Selection

Data set taken for analysis may contain thousands of attributes [14]. The entire feature set may not be important for building a classifier model. So, we pick out the subset of features which are relevant for classifier construction. The methods for feature selection include the following:

Information Gain (IG): It uses entropy [17]-based feature evaluation technique to determine the amount of information about the lexical features needed for text review classification. Entropy is the sum of the probability of each label times the log probability of that same label. The expected amount of information needed to classify the review on the training data set D of m labeled classes is defined as, $E(D) = -\sum_{l=1}^m (P_l)\log_2(P_l)$ where P_l denotes the probability that the review

belongs to the class C_l on D . P_l can be estimated as the number of reviews belongs to one particular class for the given total number of reviews. If the review in D on some feature f_i with set of values say, $\{a_1, a_2, \dots, a_v\}$, then the data set will have $\{D_1, D_2, \dots, D_v\}$ partitions set. The expected amount of information needed to classify the reviews in the training data set D , if the reviews are classified according to one particular feature f_i is, $E(D, f_i) = \sum_{j=1}^v \frac{|D_j|}{|D|} \times E(D_j)$ where $\frac{|D_j|}{|D|}$ is the weight of j th partition, and $E(D_j)$ is the entropy of a partition D_j . The information gain on feature f_i is $Information_gain(f_i) = E(D) - E(D, f_i)$. The features are ranked as per the highest information gain value [14].

Information Gain Ratio (GR): GR is an iterative process for choosing the subset of features [17, 18]. The information gain measure is biased because it is tested with the attributes having a large number of outcomes. GR overcomes the biasness by applying normalization to the IG by utilizing a measure called split information [12]. The split information measure for an attribute f_i is generated by splitting the data set D into v partitions where v is the outcome of a test on attribute f_i . The split information of a particular feature f_i is defined as $Split_information_{f_i}(D) = - \sum_{j=1}^v \frac{|D_j|}{|D|} \times \log_2 \frac{|D_j|}{|D|}$. If the split information is high, then all the partitions have a uniform number of review instances, else smaller number of partitions with more review instances. The gain ratio of an attribute f_i for data set D is defined as $Gain_Ratio_{f_i} = \frac{Information_gain(f_i)}{Split_information(f_i)}$.

CHI Square Statistic (CHI): CHI measures the association between the feature f_i and class C_l [6, 14]. If the CHI measure of the feature f_i for the class C_l is large, then the ability of the feature f_i classifying the text review is strong. The CHI value with feature f_i for binary classification is

$$\chi^2(f_i, C_l) = \frac{N \times [N(C_{1,f_i}) \times N(C_{2,\bar{f}_i}) - N(C_{2,f_i}) \times N(C_{1,\bar{f}_i})]^2}{\{[N(C_{1,f_i}) + N(C_{2,f_i})] \times [N(C_{2,\bar{f}_i}) + N(C_{1,\bar{f}_i})] \\ \times [N(C_{1,f_i}) + N(C_{1,\bar{f}_i})] \times [N(C_{2,f_i}) + N(C_{2,\bar{f}_i})]\}}$$

where N represents the total number of text reviews, $N(C_b, f_i)$ represents the number of reviews that belongs to class C_l containing feature f_i , $N(C_l, \bar{f}_i)$ represents the number of reviews that belongs to class C_l does not containing the feature f_i . The variable l takes the values 0 and 1 [6].

One Rule (OneR): The OneR feature selection method evaluates each attribute separately. It treats all the attribute value as continuous and the attribute values are split into disjoint ranges. OneR computes class distribution and produces rule for each attribute based on the each attribute value. The rule has been built by finding the most frequent class based on the attribute value. In the end, it calculates error rate of each rule of each attribute on the entire data set and rank the rule in descending order of error rate [19].

Relief-F (RF): The weight of an individual feature has been evaluated by sampling the instances repeatedly and considering the value of the given feature for the nearest instances of same and different class [14]. Each feature has been weighted to distinguish the instance, among the classes. The weight of an attribute is the difference between the probability value of the two nearest neighbors of the same class having the same value of features and the probability value of the two nearest neighbors of the different class having the different values of a feature. The feature is believed to be more significant only when the difference between the probabilities is high. The weight of the feature is calculated as

$$\begin{aligned} W[F] = & W[F] - \sum_{j=1}^k \frac{\text{diff}(F, C, H_j)}{(m * k)} \\ & + \sum_{C \neq \text{class}(I_j)} \left[\frac{P(C)}{1 - P(\text{class}(I_j))} \sum_{j=1}^k \frac{\text{diff}(F, I_j, M_j)}{(m * k)} \right] \end{aligned}$$

where $W[F]$ is feature weight, I_j is the randomly selected instances, H_j is hit when the same class having same feature value, M_j is a miss when different class having different feature values, and m is the user-defined value which represents the number of the times the process is repeated [20].

Significance Attribute Evaluation (SAE): The attribute weight has been measured by estimating the probability significance of any instances of Class-Feature and Feature-Class Association [16]. Attribute-Class association captures the cumulative value of all features and their association with the class label. Similarly, Class-Attribute association has been obtained by estimating how a feature value changes with the change in the class label. The feature is said to be significant only when both the Class-Feature and Feature-Class associations are high and those significant features are ranked accordingly. The Feature-Class Association is denoted as $P_i^r(w) + P_i^{\sim r}(\sim w)$, where w be the subset of a binary class labeled data set, $P_i^r(w)$ is the probability of frequency count of instances of w with r th feature values class equals to the i th feature values of an instance in the whole data set, and $P_i^{\sim r}(\sim w)$ is the probability of frequency count of class instances of w with r th feature values, not equals to the i th feature values of an instance in the whole data set. The Class-Feature association is denoted as $P_i^j(V) + P_i^{\sim j}(\sim V)$, where V be the subset of feature values of F_i , $P_i^j(V)$ is the probability of frequency count of an instances that are belonging to the class j contains the feature values of F_i which are same the feature value in the set V , and $P_i^{\sim j}(\sim V)$ is the probability of frequency count of an instances that are not belonging to the class j contains the feature values of F_i which are not same the feature values in the set V .

3.2 Feature Weighting

Feature weighting techniques are employed to improve the classifier accuracy where each and every feature has different capabilities to classify the review accurately. Term presence (TP) [3, 7, 13] weighting scheme is considered in this work where the value 1 represents the presence of the feature and 0 otherwise.

3.3 Classification Technique

Each classifier gives a model that fits the relationship between the feature set and class label. The classifier techniques used in this paper include the following:

Support Vector Machine: SVM is the non-probabilistic linear classifier model [2–4, 7–9, 10–13]. A SVM classifier model has been built by finding hyper plane that separates the sample reviews into two non-overlapping classes with a margin of maximum distance. The reviews, which defines the hyper plane, are called support vectors represented by vector \vec{w} . The optimized hyper plane can be written as $\vec{w} = \sum_j \alpha_j C_j \vec{r}_j$ where α_j are support vectors and found by solving a dual optimization problem which is greater than zero, C_j is a correct class category for a review vector \vec{r}_j .

If the data points cannot be separated linearly, then the kernel function can be applied which map the low-dimensional nonlinear data points into high-dimensional linear data points. In this paper, the SVM parameters are optimized using the linear kernel.

Naive Bayes: Naive Bayes is the probabilistic classifier model which uses Bayes rule [3, 8]. Consider a review r , whose class is c , and can be classified as either positive or negative based on the posterior probability value $P(c|r)$ as represented in Bayes theorem $P(c|r) = \frac{P(r|c)P(c)}{P(r)}$. In NB, each feature is considered to be independent of

each other for a given class. The NB theorem is $P(c|r) = \frac{P(c) \prod_{i=1}^m P(f_i|c)}{P(r)}$, where f_i represents the features in the text review r .

Decision Tree: C4.5 is a well-known decision tree algorithm which is based on the concept of information theory. The tree has been constructed by considering a best splitting attribute at each level identified using the information gain measure [7]. DT has been interpreted easily and represented in graphical format. The leaves in the tree represent the class label, and branches in the tree represent the condition on a particular attribute [17, 18].

K-Nearest Neighbor: K-NN is a lazy or instant-based classifier [14] that relies on the class labels of training review which are similar to the test review. The reviews

are classified accordingly to the most common among its k -nearest neighbors. The similarity measure of each nearest neighbor review to the test review r is used as weight of class of the neighbor review. The weighted sum in K-NN is $\text{Score}(r, c_l) = \sum_{r_i \in \text{KNN}(r)} \text{sim}(r, r_i) * \delta(r_i, c_l)$ where $\text{K-NN}(r)$ represents the review set of k -nearest neighbors of review r . If r_i belong to class c_l , then $\delta(r_i, c_l)$ is 1, otherwise 0. The review should belong to the class of the review that has the highest resulting weighted sum.

Maximum Entropy: The maximum entropy classifier is a probabilistic model that maximizes the likelihood which is about exponential model [14]. It estimates the conditional distribution of the class label c given the review r with features f_i by following exponential form

$$P(c/r) = \frac{1}{\sum_c \exp(\sum_i \lambda_i f_i(r, c))} \exp\left(\sum_i \lambda_i f_i(r, c)\right)$$

where $f_i(r, c)$ is a feature function obtains value 1 when $r = f_i$ and $c = c_i$, and 0 otherwise, λ_i is a weight coefficient.

3.4 Performance Measures

Accuracy measures used to judge the performance of review classification. For binary text review classification, accuracy is defined as the ratio of the number of predictions that are correct given the total reviews [17].

4 Experimental Results

4.1 Data set and Experimental Design

The benchmark data set used for sentiment classification is the binary class movie review data set known as polarity data set v2.0 with 1,000 positive and 1,000 negative reviews collected from the Internet Movie Database (IMDb) [21]. Python is used for the execution of all feature selection techniques and classifier model. Each feature represents the unigram word extracted from the review, and feature vector model is utilized for building classifier. The informative features are ranked and all the classifiers are constructed and validated using 10-fold cross-validation with the TP weighting scheme.

4.2 Comparison of Feature Selection and ML Methods

In order to evaluate the potential features for classification, the comparative study is performed on various feature selection methods with ML techniques. Table 2 shows the accuracy of the six feature selection mechanisms combined with five ML techniques. From Table 2, it is observed that NB outperforms all other classifier techniques for the feature selection methods IG, GR, CHI, and SAE with an accuracy of 92.95, 95.3, 92.95, and 95.6 %, respectively. And ME outperforms all other classifier techniques for the feature selection methods OneR and RF with an accuracy of 88.6 and 70.65 %, respectively.

NB performed better because it is based on the probability of occurrences of words in the text. Though NB performs better for a movie review data set, the range of features for which it performs better should be conceived. With respect to Figs. 1, 2, 3, 4, 5, and 6, it is mentioned that the NB, SVM, and ME gave a more respectable

Table 2 Accuracy for different classifiers (in %)

	% Feature selection mechanisms	Classifier techniques					
		IG	GR	CHI	OneR	RF	SAE
SVM	90.35	90.7	90.35	87.65	67.65	90.35	
NB	92.95	95.3	92.95	86.9	69.95	95.6	
DT	65.6	65.95	80.85	63.85	59	81.05	
K-NN	66.3	66.3	92.7	60.75	57.05	95.4	
ME	90.8	92.85	90.8	88.6	70.65	94.4	

Fig. 1 Performance of ML in IG

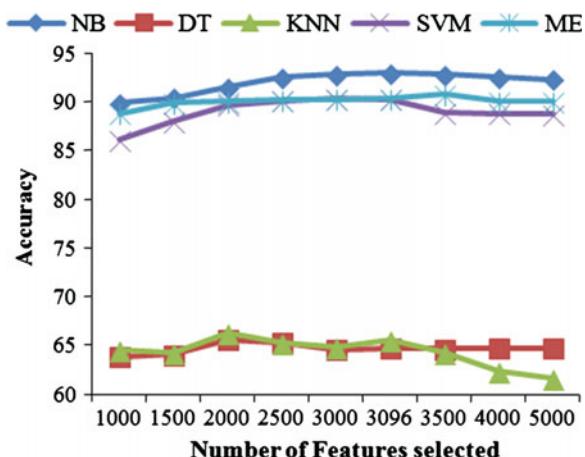


Fig. 2 Performance of ML in CHI

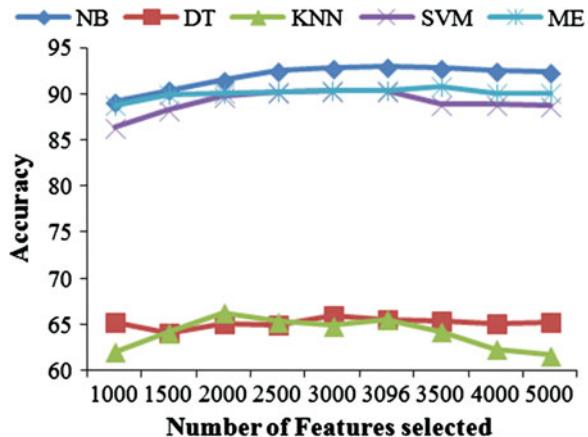
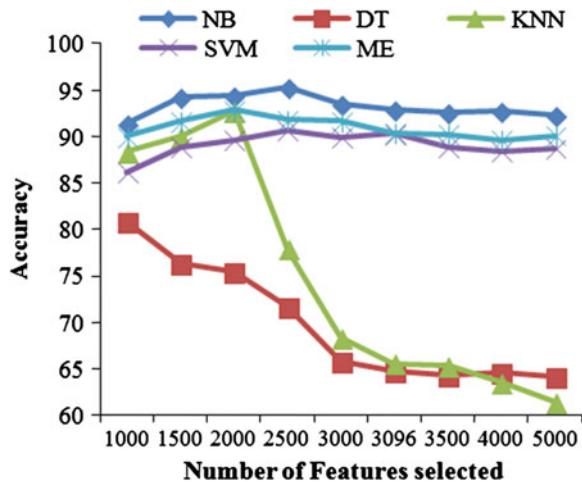


Fig. 3 Performance of ML in GR



performance in the range of 2,500–3,500 features, whereas K-NN and DT performs better for features above 2,000.

From Fig. 6, it is observed that the RF feature selection method gives poor accuracy for the movie review date set. From Fig. 7 it is informed that SAE, GR, IG, CHI, and OneR feature selection methods give better accuracy for the NB classifier. Overall, we observe that NB with SAE and GR feature selection technique performs best for Movie review data set with around 2,500 informative features. The experimental result shows that the classifier gives improved accuracy and the accuracy remains stable for the feature more than 2,500.

Fig. 4 Performance of ML in OneR

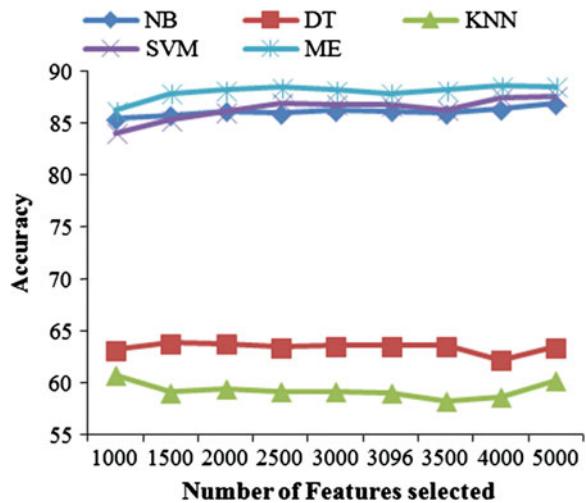


Fig. 5 Performance of ML in SAE

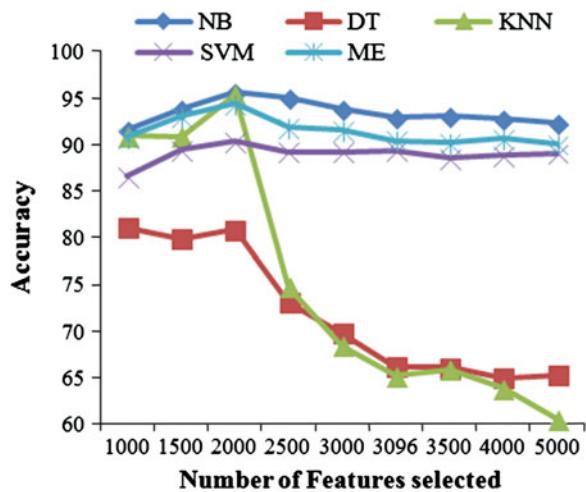


Fig. 6 Performance of ML in RF

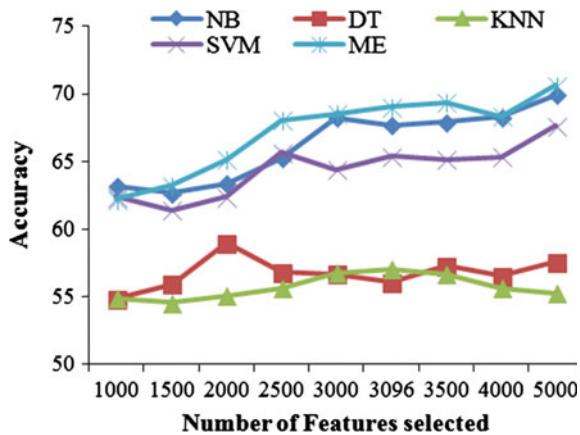
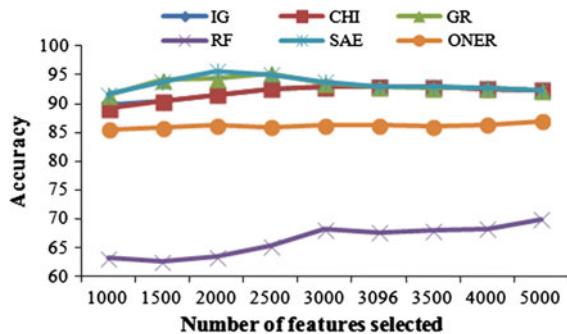


Fig. 7 Performance of NB in IG, CHI, GR, SAE, RF, and OneR



5 Conclusion

The paper explored the significance of applying different feature selection methods IG, GR, CHI, OneR, RF, and SAE on various classifiers such as NB, SVM, DT, K-NN, and ME for binary text classification and investigates the performance in terms of accuracy. We have shown that confining to the range of 2,500–3,500 features gives better results than on entire unigram set. Results show that significance attribute evaluator and GR perform better among sentiment feature selection methods and NB performs better among classifiers. Our upcoming work is to investigate the performance of feature selection methods applied to preprocessed text and POS features on different machine learning classifiers and to evaluate the model for multiple data set other than movie review data set.

References

1. Wang, G., Sun, J., Ma, J., Xu, K., Gu, J.: Sentiment classification: the contribution of ensemble learning. *Decis. Support Syst.* **57**, 77–93 (2014)
2. Boiy, E., Hens, P., Deschacht, K., Moens, M.F.: Automatic sentiment analysis in on-line text. In: ELPUB (2007)
3. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up?: sentiment classification using machine learning techniques. In: ACL-02 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, vol. 10 (2002)
4. Kennedy, A., Inkpen, D.: Sentiment classification of movie reviews using contextual valence shifters. *Comput Intell* **22**(2), 110–125 (2006)
5. Wang, S., Li, D., Song, X., Wei, Y., Li, H.: A feature selection method based on improved fisher's discriminant ratio for text sentiment classification. *Expert Syst. Appl.* **38**(7), 8696–8702 (2011)
6. Tan, S., Zhang, J.: An empirical study of sentiment analysis for Chinese documents. *Expert Syst. Appl.* **34**(4), 2622–2629 (2008)
7. Annett, M., Kondrak, G.: A comparison of sentiment analysis techniques: polarizing movie blogs. In: Advances in Artificial Intelligence, pp. 25–35. Springer, Heidelberg (2008)
8. Xia, R., Zong, C., Li, S.: Ensemble of feature sets and classification algorithms for sentiment classification. *Inf. Sci.* **181**(6), 1138–1152 (2011)
9. Moraes, R., Valiati, J.F., Gavião Neto, W.P.: Document-level sentiment classification: an empirical comparison between SVM and ANN. *Expert Syst. Appl.* **40**(2), 621–633 (2013)
10. Pang, B., Lee, L.: A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts. In: Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics (2004)
11. Pang, B., Lee, L.: Seeing stars: exploiting class relationships for sentiment categorization with respect to rating scales. In: Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics (2005)
12. Dasgupta, S., Ng, V.: Topic-wise, sentiment-wise, or otherwise?: identifying the hidden dimension for unsupervised text classification. Paper Presented at the 2009 Conference on Empirical Methods in Natural Language Processing, vol. 2. Association for Computational Linguistics (2009)
13. Paltoglou, G., Mike, T.: A study of information retrieval weighting schemes for sentiment analysis. In: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics (2010)
14. Sharma, A., Dey, S.: A comparative study of feature selection and machine learning techniques for sentiment analysis. In: Proceedings of the Symposium on Research in Applied Computation, ACM (2012)
15. Samsudin, N., Puteh, M., Hamdan, A.R., Nazri, M.Z.A.: Immune based feature selection for opinion mining. In: World Congress on Engineering, vol. 3 (2013)
16. Ahmad, A., Dey, L.: A feature selection technique for classificatory analysis. *Pattern Recogn. Lett.* **26**(1), 43–56 (2005)
17. Han, J., Kamber, M.: Data Mining, Concepts and Techniques. Southeast Asia Edition. Morgan Kaufmann, Los Altos (2006)
18. Quinlan, J.R.: C4.5: Programs for Machine Learning, vol. 1. Morgan Kaufmann, Los Altos (1993)
19. Morariu, D.I., Crețulescu, R.G., Breazu, M.: Feature selection in document classification
20. Robnik-Šikonja, M., Kononenko, I.: Theoretical and empirical analysis of ReliefF and RReliefF. *Mach. Learn.* **53**(1–2), 23–69 (2003)
21. Movie Review Data. <http://www.cs.cornell.edu/people/pabo/movie-review-data>

Enhancement of Transmission System Loadability During Contingency by Optimal Allocation of FACTS Devices Using Particle Swarm Optimization

P. Malathy, A. Shunmugalatha and P. Thaineesh

Abstract This paper deals with the enhancement of transmission system loadability with single contingency by optimal location and settings of multi-type Flexible AC Transmission System (FACTS) devices such as Thyristor-Controlled Series Compensator (TCSC) and Static Var Compensator (SVC). Contingency Severity Index (CSI) is used for the optimal location of FACTS devices. To enhance the socioeconomical benefits, the solution to the problem is optimized using Particle Swarm Optimization (PSO) technique thereby minimizing the Installation Cost (IC) of FACTS devices and Severity of OverLoading (SOL).

Keywords Maximum Loadability · Contingency · TCSC · SVC · PSO

1 Introduction

Power system stability is a major issue under contingency conditions in restructured environment [1]. To enhance the stability, contingency screening and ranking [2] is used as a tool for online analysis of power system security. FACTS devices maximize the power flow, reduce the losses, and enhance the stability of the power system network [3, 4]. Thyristor-Controlled Series Capacitor (TCSC) offers faster and smooth response under contingency conditions [5]. Static var compensator (SVC) is capable of maintaining voltage within the operating limits during normal and network contingency conditions [6]. Particle Swarm Optimization (PSO) has been used to

P. Malathy (✉) · P. Thaineesh

PSNA College of Engineering and Technology, Dindigul, Tamil Nadu, India
e-mail: malathypaulpandi@gmail.com

P. Thaineesh

e-mail: thaineesh29@gmail.com

A. Shunmugalatha

Velammal College of Engineering and Technology, Madurai, Tamil Nadu, India
e-mail: shunmugalatha@gmail.com

solve complex problems such as voltage stability [7, 8]. Contingency analysis based on contingency severity index (CSI) with the incorporation of SVC is reported in [9]. Contingency analysis [10] for power system planning has been discussed. Different computational techniques [11] for solving various power system problems have been discussed. Various aspects of voltage stability and methods of enhancing system stability and methods to prevent the voltage collapsed are reported in [12]. MATPOWER [13] is a MATLAB-based software tool to analyze various power system problems. The work proposed is to improve the voltage stability through maximizing the loadability. The maximum loadability (ML) can be increased by placing multi-type FACTS devices such as SVC and TCSC. Optimal location of TCSC and SVC is obtained to reduce the installation cost (IC) of the devices. To validate the proposed technique, WSCC-9 bus system is simulated using MATLAB. To begin with, the base case ML is obtained using continuation power flow (CPF) technique. The system is then simulated with line outage contingency, and its branches are ranked based on CSI. The worst contingency whose impact is very severe is ranked first. The placement of FACTS devices is implemented using binary decision. The numbers and settings of these devices are optimized using PSO to maximize the loadability, to improve the voltage stability, and to minimize the cost of installation of FACTS devices thereby achieving both social and economical benefits.

2 Methodology

2.1 Problem Formulation

The objective function to maximize the loading factor λ and to minimize the cost of installation of FACTS devices is given by

$$\begin{aligned} F &= \text{Max } Z + \text{Min IC} \\ Z &= \lambda \\ \text{IC} &= S_{ab} * C_d * 1,000 \end{aligned} \tag{1}$$

where

λ Loading factor

IC Optimal cost of installation of multi-type FACTS devices

C_d Cost of installation of FACTS device(s)

d FACTS device(s) SVC and/or TCSC

S_{ab} $Q_a - Q_b$, VAr compensation provided by SVC and TCSC in KVAr

Q_b MVAr flow in the branch before placing the FACTS devices

Q_a MVAr flow in the branch after placing the FACTS devices

This objective function is subjected to the equality and inequality constraints. The equality constraints include voltage stability and power balance constraints

given by Eqs. (2)–(5). The constraints for FACTS devices contribute for inequality given by Eqs. (6) and (7).

$$\begin{aligned} V_s &= 1 \quad \text{if } V_{\min} \leq V_i \leq V_{\max} \\ &= 0 \quad \text{otherwise} \end{aligned} \quad (2)$$

$$\sum_{i=1}^{n_b} P_{gi} = \sum_{i=1}^{n_b} P_{di} + P_{\text{loss}} \quad (3)$$

$$P_i = \sum_{i=1}^{n_b} |E_i| |E_j| [G_{ij} \cos(\theta_i - \theta_j) + B_{ij} \sin(\theta_i - \theta_j)] \quad (4)$$

$$Q_i = \sum_{i=1}^{n_b} |E_i| |E_j| [G_{ij} \sin(\theta_i - \theta_j) + B_{ij} \cos(\theta_i - \theta_j)] \quad (5)$$

$$-0.5 X_{\text{Line}} \leq X_{\text{TCSC}} \leq 0.5 X_{\text{Line}} \quad (6)$$

$$-200 \text{ MVAr} \leq Q_{\text{SVC}} \leq 200 \text{ MVAr} \quad (7)$$

where

P_{gi} , P_{di}	Real power generation and real power demand, respectively
P_{loss}	Transmission line losses
θ_i , θ_j	Phase angles at buses i and j , respectively
E_i , E_j	Voltage magnitudes at buses i and j , respectively
G_{ij} , B_{ij}	Conductance and susceptance of the Y bus matrix
V_i	Voltage at bus i
X_{Line}	Actual line reactance
X_{TCSC}	Reactance offered by TCSC
Q_{SVC}	Var compensation offered by SVC
P_i , Q_i	Real power and reactive power at bus i
V_{\min} , V_{\max}	Minimum and maximum voltage limit of bus i
n_b	Number of buses

2.2 Optimal Location of FACTS Devices

The step-by-step procedure for optimal placement of TCSC and SVC is given below.

Step 1 Find the participation matrix. The participation matrix (U) is used for identifying the branches which are overloaded. The size of the binary matrix ($x * y$) consists of 1's or 0's based on whether or not the corresponding branch is overloaded, where q is the total number of branches of interest and p is the total number of single contingencies.

Step 2 Find the ratio matrix. The elements of ratio matrix (W) are normalized overflow of all the branches. The size of the matrix is ($x * y$) in which each element denotes the excess power flow with respect to the base case flow through branch y during contingency x .

$$W_{xy} = \frac{P_{xy, \text{Conti}}}{P_{oy, \text{Base}}} - 1 \quad (8)$$

where

$P_{xy, \text{Conti}}$ Power flow through branch y during contingency x
 $P_{oy, \text{Base}}$ Base case power flow through branch y

Step 3 Find the contingency probability array (P). This array is used to identify the probability of branch outages. It is a ($p * 1$) array. The probability of branch outage is calculated based on the equation given below.

$$P_{p \times 1} = [P_1 P_2 \dots P_p] \quad (9)$$

where

$P_{xy, \text{Conti}}$ Power flow through branch y during contingency x
 $P_{oy, \text{Base}}$ Base case power flow through branch y

Step 4 Find the CSI for all the branches. The CSI is evaluated for each and every branch considering all possible single contingencies using Eq. (10), and the best location for the placement of FACTS devices is identified.

$$\text{CSI}_y = \sum_{x=1}^p P_x U_{xy} W_{xy} \quad (10)$$

where

CSI_y Contingency severity for the branch y
 P_x Elements of contingency probability array
 U_{xy} Elements of participation matrix U
 W_{xy} Elements of ratio matrix W

2.3 Optimal Settings of FACTS Devices

In order to minimize the severity of overloading given by Eq. (10), the best possible setting of FACTS devices is determined for all possible single contingencies, and the optimization problem will have to be solved using PSO technique.

$$\text{SOL} = \sum_{c=1}^p \sum_{y=1}^q [a_y] \left[\frac{P_y}{P_{\text{ymax}}} \right]^4 \quad (11)$$

where

a_y	Weight factor
P_y	Real power flow on branch k
P_{ymax}	Maximum real power flow
p	Number of contingencies considered
q	Number of branches
c	Contingency index

2.4 Installation Cost

The cost of TCSC and SVC is given by Eqs. (12) and (13),

$$C_{\text{TCSC}} = 0.0001S_{ab}^2 - 0.71S_{ab} + 153.75(\text{US\$/KVar}) \quad (12)$$

$$C_{\text{SVC}} = 0.0003S_{ab}^2 - 0.305S_{ab} + 127.38(\text{US\$/KVar}) \quad (13)$$

where $S_{ab} = Q_a - Q_b$, is the reactive power compensation is provided by SVC and TCSC in KVar. Q_a and Q_b are the KVar flow through the branch after and before placing the FACTS devices respectively.

3 Particle Swarm Optimization

PSO is an evolutionary computational technique which is based on the simulation of bird flocking. PSO consists of a search space with population of possible solutions for the given problem. The position of each agent is represented in x - y plane as (X_x, X_y) , moving with velocities (V_x, V_y) along x -axis and y -axis. The position of each agent is modified by updating its velocity. Each agent knows its best value so far, called local best, ' P_{besti} '. It contains the information on positions and velocities. Each agent knows the best value in the group, called global best, ' G_{besti} ', among all P_{best} . Each agent modifies its position by considering current positions (X_x, X_y) , velocities (V_x, V_y) , the individual intelligence (P_{besti}), and the group intelligence (G_{besti}). The velocity, position, and inertia weight are updated using Eqs. (14)–(16).

$$V_i^{j+1} = W * V_i^j + C_1 * R_1 * (P_{\text{besti}} - X_i^j) + C_2 * R_2 * (G_{\text{besti}} - X_i^j) \quad (14)$$

$$X_i^{j+1} = X_i^j + V_i^j \quad (15)$$

$$W = W_{\max} - \frac{(W_{\max} - W_{\min})}{\text{iter}_{\max}} * \text{iter} \quad (16)$$

where

V_i^{j+1}	Velocity of i th individual at $(j + 1)$ th iteration
V_i^j	Velocity of i th individual at j th iteration
W	Inertia weight
C_1, C_2	Learning factors both equal to 2
R_1, R_2	Random numbers selected between 0 and 1
P_{besti}	Local best of i th individual
G_{besti}	Global best of i th individual
X_i^j	Position of the i th individual at j th iteration
W_{\max}	Maximum value of inertia weight
W_{\min}	Minimum value of inertia weight
iter	Current iteration number
iter _{max}	Maximum number of iterations

4 Results and Discussion

This section presents a case study based on WSCC-9 bus system, in order to validate the PSO technique for ML problem by considering single contingency. For this case study, five number of FACTS devices have been considered. The problem is optimized using PSO with 50 different initial solutions with variables as load demand and reactance and reactive power. The program is done using MATLAB 4.1 and simulated by using Intel (R) core (TM) i3-2330 M processor with a speed of 2.2 GHz. The maximum number of iterations and runs considered for PSO is limited to 50. The numerical results for the simulation are obtained and tabulated

Table 1 CSI-based ranking for branches

Rank	Branch	CSI
1	6–7	0.6065
2	8–9	0.5337
3	7–8	0.5273
4	4–5	0.4879
5	8–2	0.4798
6	9–4	0.4308
7	3–6	0.3531
8	5–6	0.3531
9	1–4	0.1458

from Tables 1, 2, 3, and 4. The ranking of branches depending upon the CSI for all possible single contingencies is shown in Table 1.

From Table 1, it is inferred that the branch (6–7), which has high severity index, is given first priority for the placement of FACTS devices. Then, preferences are given to other branches such as (8–9), (7–8), (4–5), and so on. The type, size, and cost of FACTS devices are optimized using PSO algorithm. The ML obtained for base case without any FACTS devices is 35.12 %.

The ML without considering contingency with FACTS devices using PSO is shown in Table 2. It shows that the ML increases with more number of FACTS devices. The ML without considering contingency and optimized using PSO is 35.15 %.

The system is simulated for line outage. The ML obtained during line contingency with FACTS devices using PSO is shown in Table 3. From this table, it is clear that ML increases with more number of FACTS devices. The ML optimized using PSO is 34.15 % by placing four FACTS devices, namely three SVCs at buses 7, 8, and 9 and one TCSC at branch 5.

Table 4 shows the individual and total cost of FACTS devices placed in the system and its corresponding locations and settings of TCSC and SVC, namely reactance and reactive power to obtain the ML and hence to enhance the system stability using PSO, without considering contingency. It is inferred that by placing one SVC at bus 9, ML obtained is 35.11 % at 40th run with the convergence time of 0.7117 s. By placing two SVCs at buses 8 and 9, ML obtained is 35.12 % at 29th run with the convergence time of 0.7463 s. By placing two TCSCs at branches 5 and 8 and one SVC at bus 9, ML obtained is 35.14 % at 46th run with the convergence time of 1.4921 s. By placing one TCSC at branch 5(6–7) and three SVCs at buses 7, 8, and 9, ML obtained is 35.14 % at 31st run with the convergence time of 1.9064 s. By placing two TCSCs at branches 5 and 8 and three SVCs at buses 7, 8, and 9, ML obtained is 35.15 % at 46th run with the convergence time of 0.7411 s. It is also evident from the table that the ML

Table 2 Optimized ML without contingency with FACTS devices using PSO

No. of devices	Placement of devices		ML (%)
	Bus no. (SVC)	Line no. (TCSC)	
1	9	–	35.11
2	8, 9	–	35.12
3	9	5, 8	35.14
4	7, 8, 9	5	35.14
5	7, 8, 9	5, 8	35.15

Table 3 Optimized ML considering single line contingency with FACTS devices using PSO

No. of devices	Placement of devices		ML (%)
	Bus no. (SVC)	Line no. (TCSC)	
1	–	5	33.08
2	9	5	34.13
3	9	5, 8	33.97
4	7, 8, 9	5	34.15
5	7, 8, 9	5, 8	34.06

Table 4 Optimized ML, cost, and settings of FACTS devices without contingency using PSO

No. of devices	ML (%)	Time of convergence (s)	Location	Cost of FACTS devices (\$/KVAr)			Reactive power Q (MVAr)
				TCSC	SVC	Total	
1	35.11	0.7117	9	—	127.3800	127.3800	-0.3742
	35.12	0.7463	8	—	127.3800	254.7600	29.2032
			9	—	127.3800	—	79.2032
3	35.14	1.4921	5 (6-7)	153.6161	—	434.9430	0.1128
			8 (8-9)	154.2100	—	—	0.1389
			9	—	127.1168	—	15.5241
4	35.14	1.9064	5 (6-7)	154.7566	—	538.6421	0.1360
			7	—	128.7913	—	-72.5671
			8	—	127.6931	—	-107.567
5	35.15	0.7411	9	—	127.4011	—	-57.5671
			5 (6-7)	154.1748	—	691.3324	0.1311
			8 (8-9)	153.0805	—	0.2094	—
			7	—	129.5925	—	113.4956
			8	—	128.0029	—	128.4956
			9	—	126.5817	—	78.4956

Table 5 Optimized ML, cost and settings of FACTS devices with single line contingency using PSO

No. of devices	ML (%)	Time of convergence (s)	Location	Cost of FACTS Devices (\$/KVAR)			Reactance X (p.u)	Reactive Power Q (MVAr)
				TCSC	SVC	Total		
1	33.08	0.8614	5	153.8759	—	153.8759	0.0789	—
2	34.13	0.5506	5 (6-7)	153.9023	—	281.4862	-0.0776	—
3	33.97	0.8838	9	—	127.5838	—	32.8154	—
4	34.15	2.8179	5 (6-7)	154.7172	—	538.5350	0.1344	—
			7	—	128.7360	—	-150.307	—
			8	—	127.4009	—	-185.307	—
			9	—	127.4009	—	-135.307	—
5	34.06	1.1316	5 (6-7)	153.3126	—	686.3721	0.0753	—
			8 (8-9)	153.1446	—	0.0567	—	—
			7	—	125.9687	—	18.3188	—
			8	—	127.1220	—	-16.6812	—
			9	—	126.8443	—	33.3188	—

increases with more number of FACTS devices. Furthermore, the system behavior is analyzed for single line outage contingency.

Table 5 shows the optimal location, settings, and cost of FACTS devices to obtain the ML, using PSO, by considering line contingency. It is inferred that by placing one TCSC at branch 5, ML obtained is 33.08 % at 17th run with the convergence time of 0.8164 s. By placing one SVC at bus 9 and one TCSC at branch 5, ML obtained is 34.13 % at 25th run with the convergence time of 0.5506 s. By placing two TCSCs at branch (6–7) and (8–9) and one SVC at bus 9, ML obtained is 33.97 % at 37th run with the convergence time of 0.8838 s. By placing one TCSC at 5th line that branches between buses 6 and 7 and three SVCs at buses 7, 8, and 9, ML obtained is 34.15 %.

Fig. 1 Number of FACTS devices versus ML obtained with contingency using PSO

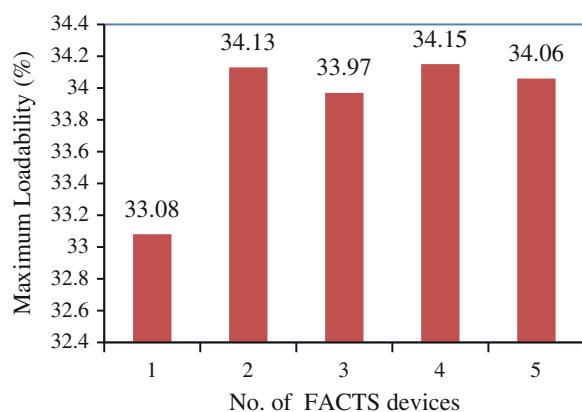
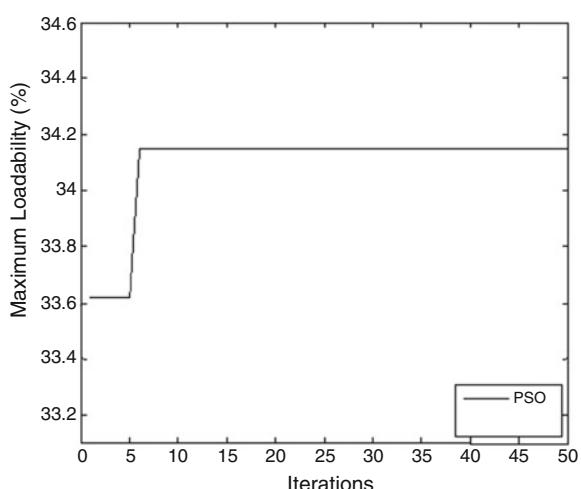


Fig. 2 Convergence characteristics obtained with line contingency using PSO



In the bar chart shown in Fig. 1, ML is obtained by considering line contingency using PSO. The ML of 34.15 % is obtained with four numbers of FACTS devices using PSO.

Figure 2 shows the convergence characteristics obtained by considering line contingency using PSO. The ML of 34.15 % is obtained with only four numbers of FACTS devices using PSO.

5 Conclusion

This work presents, the PSO technique to enhance the voltage stability by maximizing the loadability of a transmission system through the optimal placement of TCSC and SVC by considering all possible single contingencies including generators and branches. The proposed algorithm has proved its efficiency in obtaining optimal or near-optimal solutions. The case study analyzed proves the same. To demonstrate the effectiveness and robustness of the proposed technique, Newton–Raphson power flow method is used for analyzing the voltage stability improvement by increasing the ML through the placement of series and shunt-type FACTS devices, namely TCSC and SVC. This method shows the optimal location for placement of TCSC and SVC devices and their optimal settings for the improvement of ML by installing devices economically. The solutions obtained are superior to those obtained using the conventional power flow method. This work can be extended for the enhancement of ML by using other FACTS devices such as Unified Power Flow Controller (UPFC) and Thyristor-Controlled Phase Angle Regulator (TCPAR), and in future, this work will be proceeded by considering multiple contingencies and can be optimized by using new efficient optimization algorithms.

References

1. Vinu Priya, G., Jerril Gilda, S.: Enhancement of available transfer capability using facts devices in deregulated electricity market. *Int. J. Eng. Technol. (IJETI)* **1**(2) (2014). ISSN (Online) 2348-0866
2. Sakthivel, S., Mary, D.: Reactive power optimization for voltage stability limit improvement incorporating TCSC device through DE/PSO under contingency condition. *IU-JEEE* **12**(1), 1419–1430 (2012)
3. Singh, B., Sharma, N.K., Tiwari, A.N.: A study on enhancement of loadability of large-scale emerging power systems by using FACTS controllers. *Int. J. Comput. Sci. Eng. (IJCSE)* **02** (05):1893–1903 (2010)
4. Bhedsadiya, R.H., Patel, C.R., Patel, R.M.: Transmission line loadability improvement using FACTS devices. *Int. J. Res. Eng. Technol. (IJRET)*. ISSN 2319-1163, ISSN: 2321-7308
5. Vanitha, R., Baskaran, J., Sudhakaran, M.: Application of de based WAFGP in multi objective optimal power flow using TCSC. *Indian Streams Res. J.* **4**(5) (2014). ISSN 2230-7850

6. Manikandan, B.V., Charles Raja, S., Venkatesh, P.: Available transfer capability enhancement with facts devices in the deregulated electricity market. *J. Electr. Eng. Technol.* **6**(1), 14–24 (2011)
7. Bhattacharyya, B., Goswami, S.K., Gupta, V.K.: Particle swarm intelligence based allocation of FACTS controller for the increased load ability of power system. *Int. J. Electr. Eng. Inform* **4**(4) (2012)
8. Saini, P., Sanjay K.J.: A particle swarm optimization for maximum loadability. *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.* **2**(6) (2013)
9. Subramani, C., Dash, S.S., Kumar, V., Kiran, H.: Implementation of line stability index for contingency analysis and screening in power systems. *India J. Comput. Sci.* **8**(4), 585–590 (2012). ISSN 1549-3636
10. Fonteneau-Belmudes, F.: Identification of dangerous contingencies for large scale power system security assessment. Ph.D. dissertation, Department of Electrical Engineering and Computer Science, University of Liege, Belgium (2011)
11. Nagalakshmi, S., Kamaraj, N.: Secured loadability enhancement with TCSC in transmission system using computational intelligence techniques for pool and hybrid model. *Appl. Soft Comput.* **11**, 4748–4756 (2011)
12. Nagalakshmi, S., Kamaraj, N.: Comparison of computational intelligence algorithms for loadability enhancement of restructured power system with FACTS devices. *Swarm Evol. Comput.* **5**, 17–27 (2012)
13. Zimmerman, R.D., Murillo-Sánchez, C.E., Thomas, R.J.: MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Trans. Power Syst.* **26**(1), 12–19 (2011)

Fault Diagnosis of Broken Rotor Bars in Induction Motor Using Multiscale Entropy and Backpropagation Neural Network

Alok Verma and Somnath Sarangi

Abstract Interruptions in any process industry due to machinery problem induce a serious financial loss. And as we know that induction motors occupy a major area in machinery and process industry, detection of faults beforehand is a key to avoid the state of financial or production crisis in future. The present work proposes a novel algorithm for the detection of broken rotor bars in induction motor. Stator current in addition to rotor vibration in an induction motor was measured and employed for fault detection of broken rotor bar. Multiscale entropy (MSE) is used as statistic-based approach in order to tackle the nonlinear behavior existing in rotor bar using vibration and current as the diagnostic media, as both cumulatively considered describe the regularity in the diagnostic information. The proposed work presents an approach to analyze features that distinguish the rotor vibration and stator current samples of normal induction motor from those of the broken rotor bar. Further, backpropagation neural network classifier is applied over the resultant feature set which distinguishes the faulty data set from the healthy with an accuracy level of 15.5 % for vibration and 14 % for current.

Keywords Fault diagnosis • Broken rotor bar • Multiscale entropy • Neural network

1 Introduction

Induction motors play a vital role in the safe and efficient running of any industry. Early detection of faults in the motor would help to avoid costly breakdowns in industrial plant. As we know, there may exist many types of fault in induction

A. Verma (✉)

Department of Electrical Engineering, Indian Institute of Technology, Patna, India
e-mail: alokverma@iitp.ac.in

S. Sarangi

Department of Mechanical Engineering, Indian Institute of Technology, Patna, India
e-mail: somsara@iitp.ac.in

motors [1, 2], in which broken rotor bar is one of the important faults. Some of the mere consequences of broken rotor bar are poor starting performances, excessive vibrations, higher thermal stress, or torque fluctuation, etc. Various methods have been already used to detect a broken rotor bar fault. Most well-known approach for the detection of broken rotor bars in an induction motor is based on the stator current monitoring [3, 4], and some other literatures used input power and the estimated torque as diagnostic media [5–7].

Due to variations in parameters such as friction and damping, mechanical systems often exhibit nonlinear behaviors; therefore, it may change the normal current and vibration signals to complex and nonlinear [8] form. However, commonly used signal processing techniques including time- and frequency-domain techniques and advanced signal processing techniques such as time-frequency-domain and wavelet transform may have certain limitations [8]. This results in need of techniques for nonlinear dynamic parameter estimation that can provide a good alternative to extract the defect-related features hidden in the complex as well as nonlinear vibration and current samples [8]. Various experiments [9] on nonlinear dynamic parameters used for feature extraction and fault diagnosis have already been performed, and in the several literatures, approximate entropy (ApEn) was explained and selected as a working tool for fault detection. However, ApEn has found its way in the fields of physiological signal as well as vibration signal processing of rotating machine [8, 9]. However, ApEn reported more similarity in the time series, and self-matching makes ApEn mainly dependent on the length of time series [10]. To overcome these limitations of ApEn, Richman and Moorman [11] proposed a new kind of entropy, called sample entropy (SampEn), which has attracted a lot of attention [11] in this field of work which excludes self-matches.

This paper investigates the broken rotor bar fault by using stator current in addition to rotor vibration, as diagnostic media. Multiscale entropy (MSE) was used in order to tackle the nonlinear behavior existing in rotor bar using vibration and current to describe the regularity in the diagnostic information. However, backpropagation neural network was also used for the classification and detection of broken rotor bar fault. In the recent literature [12], new entropy for measure of complexity was discussed, which is called as MSE. Many researchers applied this new method to distinguish between young healthy hearts and congestive heart failure of a person [12, 13]. However, a rotating machine that is composed of mechanical components [14] which further comprises of numbers of machine complexity considering nonlinear dynamic parameters applied on single scale (Approximate entropy and sample entropy of original time series) might prove to be insufficient for characterizing machine vibration and current signals. This shortcoming inspired the implementation of multiscale entropy method with expectations of improving performances of machine fault diagnosis. To the best of the literature survey in the field of fault detection of machine, few works have been done where MSE has been used and that to only for the vibration signals. Zhang et al. [15] presented multiscale entropy (MSE), taking into account in multiple scales, was introduced as feature extraction parameter for diagnosis of fault using vibration signal. MSE with support vector machine (SVM) constitutes the proposed intelligent fault diagnosis method. Lin et al. [16] describe an

approach to distinguish the vibration signals of aligned motor from those of a misaligned one. This described feature was obtained using wavelet transform from the difference of MSE of a signal, before and after the signal is denoised. Above-said experimental result shows that classifiers based on these features obtain good and more accurate rates than those based on frequency-related features. Zhang et al. [17] explained a bearing fault diagnosis method based on MSE and adaptive neurofuzzy inference system (ANFIS), in order to tackle the nonlinearity existing in bearing vibration as well as the uncertainty inherent in the diagnostic information.

Schoen et al. [18] explained the roller bearing damage by using motor current signature analysis. They used FFT of motor current signals for successful detection of bearing damage. Therefore, in this work, authors investigate the broken rotor bar fault by using detection media like stator current as well as rotor vibration with the MSE, and that will be a new approach, in order to tackle the nonlinearity existing in vibration and current signal of broken rotor bars. So far, induced voltages in the stator due to rotor flux have generally been used to detect broken rotor bar faults. This method avoided the presence of nonlinearity like machine magnetizing characteristics due to saturation. Several spectral analysis methods have also been employed for detection of motor faults. Apart from these methods that do not take the motor nonlinearity into consideration and other spectral analysis methods for vibration and current signals from faulty motors, a new approach of multiscale entropy analysis is being proposed for extraction of information from rotor vibration and stator current that distinguishes a fault of broken rotor bars from a healthy one.

2 Experimental Analysis

2.1 Experimental Setup

Experiment is conducted on the system as shown in Fig. 1. The entire experimental setup consists of four major subsystems: three-phase induction motor, data acquisition system, various sensors, and a computer storage and display, as shown in Fig. 1. Three-phase, 1 HP induction motor without any load was used for the experiment as shown in Fig. 1a. The induction motor with a wiring enclosure is attached on the left side of the system arrangement and that wiring enclosure of the motor support permits access to the three-phase power supply and motor supply leads and this enclosure is used sufficiently to install current probe sensors on the stator wiring that supports motor current signature analysis. Figure 1b shows the wiring enclosure attached with current probe.

Two important sensors that are used for the experiment are triaxial industrial accelerometer and current probe. An accelerometer was installed above the induction motor to collect vibration signals as shown in Fig. 1d. Figure 2 shows the internal view of induction motor with broken rotor bar. The complete specification of all the components used in the experimental setup is given in Table 1.

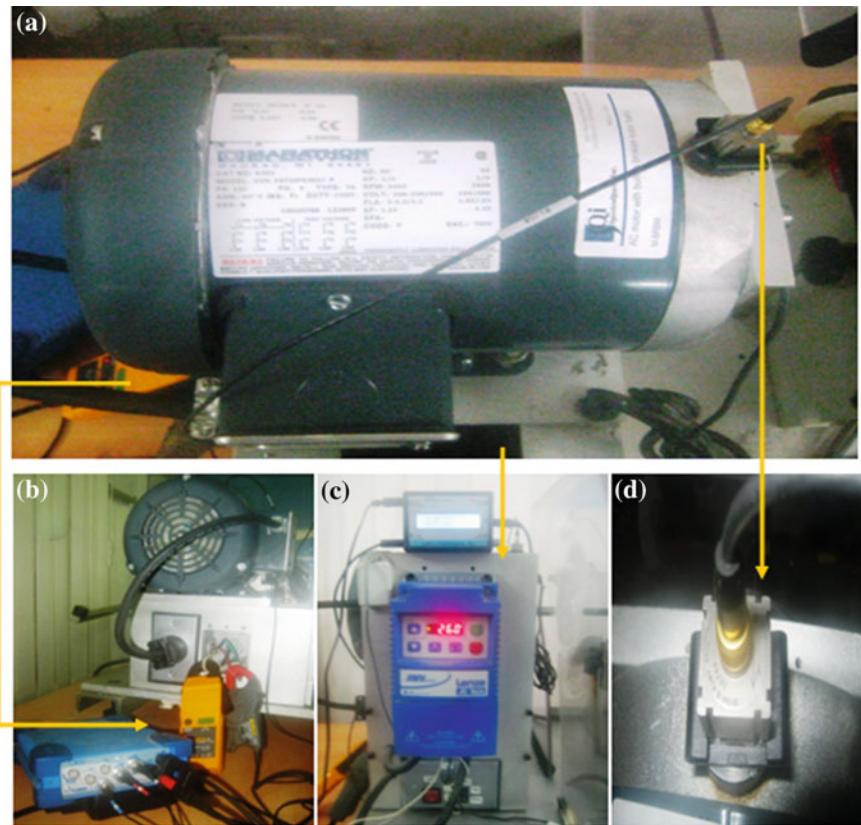


Fig. 1 Experimental setup used: **a** Three-phase induction motor. **b** Current probes and data acquisition system. **c** Frequency vector-based speed controller. **d** Triaxial industrial accelerometer

2.2 Experimental Details

The experiments were conducted with the experimental setup, which is shown in Fig. 2. In this experiment, four sets of vibration as well as current signals were collected. The first set of normal vibration and current data (denoted by N_1) was collected with the motor running in normal condition. Motor was then interchanged by normal motor to motor with broken rotor bars, and the second set data that is a faulty motor (motor with broken rotor bars denoted by F_1) was collected. Subsequently, the motor was changed back from faulty motor to normal motor, and the third set of normal data (denoted by N_2) was stored. Finally, the motor was changed again from normal to faulty one to collect the fourth set of faulty data (denoted by F_2). Each set of data contained 31 records, recording the vibration and current signals at 31 different speeds (from 700 to 1,600 rpm with an increment of 30 rpm). Each record was a time series and contains 15,364 signal values in 3 s. Vibration



Fig. 2 3- φ induction motor with broken rotor bar

Table 1 Specifications of the components used in experimental setup

Sensors/machine	Manufacturer	Model/serial no.	Sensitivity/specification
Triaxial accelerometer	Brue and Kjaer	4506	Sensitivity = 9.812 mV/(m/s ²)
Current probe	Fluke	I200s	100 mV/A
Induction motor	Marathon electric (healthy and broken rotor bars)	HVN 56T334F53033 (for healthy), HVN 56T34F5301J (for faulty)	1 HP, 50 Hz, 2,850 RPM

Table 2 Factors and levels considered for experiments

	Control factor	Level
A	Rotational speed	700–1,600 rpm
B	Motor condition	Normal and faulty

signal was taken by the accelerometer, and current signals were collected by current probe at 5.12 s/s. The parameters used in the experiment and their levels are shown in Table 2.

3 Proposed Method

The methodology used in the present investigation consists of four steps. The vibration and the current samples obtained from experiments are analyzed using MSE. Further, the results of the MSE are fed to backpropagation neural network

(BPNN) for detecting broken rotor bar of induction motor. Based on MSE of vibration and current signals, BPNN is employed to model the entire system and to give the performance. The result obtained from the above analysis is validated by performing the confirmation experiments. The proposed MSE method for the analysis is mentioned in Sect. 3.1.

3.1 Multiscale Entropy

In the present work, the MSE algorithm is used, which is based on the application of SampEn for different scales. In 1991, Pincus [10] proposed a statistical method for noisy time series, known as ApEn, to quantify the regularity of time series. In this study, the MSE algorithm is used, which is based on the SampEn for different scales of the same process instead of previously used regularity measure ApEn statistics. SampEn is used as refinement of ApEn, and it measures the regularity in series data.

ApEn (m, r, n) may be calculated for the time series of length N , according to the equation

$$\text{ApEn}(m, r, N) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left(-\ln \frac{n_i^{m+1}}{n_i^m} \right), \quad (1)$$

where r is the tolerance of time series, m is the pattern length, and n is the number of matching (including self-match). ApEn is reported with certain disadvantages, such as heavy dependence on the length of the time series, smaller values for the shorter time series, and lack of consistent results for different values of m and r . To overcome the shortcomings of ApEn, SampEn as a new kind of entropy, this excludes the disadvantages of ApEn. As a refinement of ApEn, SampEn is used and it measures the regularity in series data. The SampEn is defined as follows:

$$\text{SampEn}(m, r, N) = -\ln \left(\frac{\sum_{i=1}^{N-m} n_i^{m+1}}{\sum_{i=1}^{N-m} n_i^m} \right). \quad (2)$$

The value of r is taken as 0.15 times the standard deviation, and m is taken as 2 for the present analysis to avoid distortion. Costa et al. [13] proposed an algorithm called MSE, which considers SampEn at multiscale. MSE has been already successful to analyze physiological signals. Consider a time series, $\{X_1, \dots, X_N\}$, which is a coarse-grained time series in MSE which is given by

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} X_i \quad (3)$$

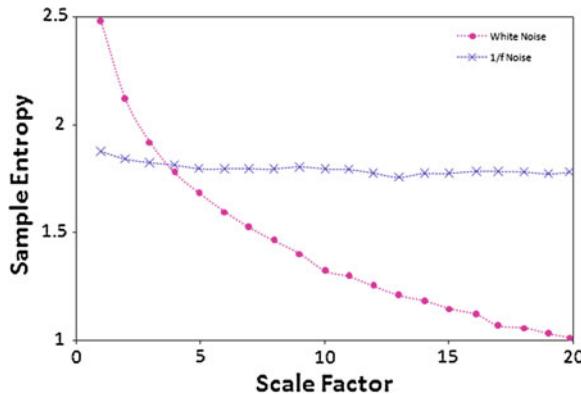


Fig. 3 Sample entropy with respect to the scale factor for coarse-grained time series of white and 1/f noises

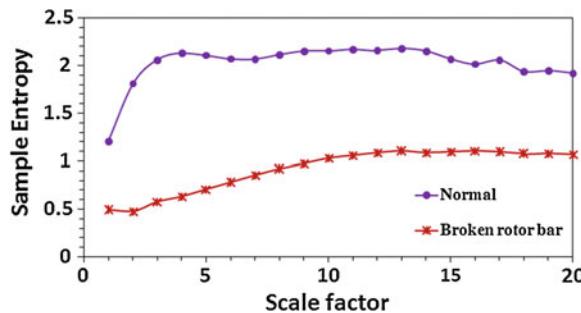


Fig. 4 Sample entropy at different scale factors of normal and broken rotor bar motor at 1,150 rpm for vibration signals

where τ is known as scale factor, and for $\tau = 1$, that coarse-grained time series is the original series. It is found that as τ increases, the length of the coarse-grained time series decreases. SampEn of white noise falls quickly as the scale factor increases as shown in Fig. 3. However, the SampEn of a pink or noise remains approximately stable as the scale factor rises. Therefore, white noise is more regular than pink noise.

Sample entropy at different scale factors of normal and faulty motors at 1,150 rpm for vibration and current signals is enunciated using Figs. 4 and 5, respectively. It can be observed through Fig. 5 that there exists increment of SampEn when the scale factor is small and reduction in SampEn when the scale factor is large. Similarly, the nature of periodicity of the current signals can also be observed from Fig. 5. Thus, the information or the distinguishable characteristic present in both vibration and current signals is extracted or inferred from its respective sample entropies and can be consequently deployed to distinguish broken rotor bar motors from that of the healthy ones.

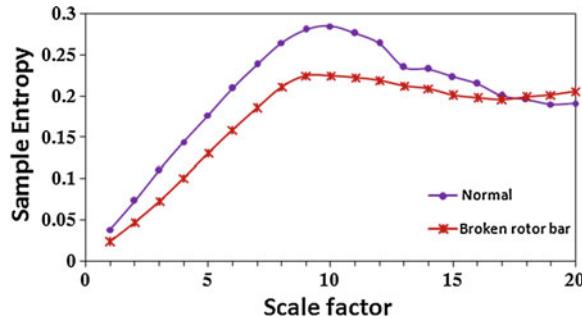


Fig. 5 Sample entropy at different scale factors of normal and broken rotor bar motor at 1,150 rpm for current signals

The sample entropy of vibration and current signals is calculated for a scale factor ranging from 1 to 20. Weighted average values of MSE are used as input to a BPNN-based classification algorithm to detect fault. Based on MSE, the performance of the classifiers is discussed in Sect. 4

4 Results and Discussion

The proposed method for fault detection of broken rotor bar underlies on the concept of statistical analysis. MSE was used for statistic-based approach in order to deal with the nonlinear behavior existing in the motor vibration and current which describes the regularity in the diagnostic information. The present work shows an approach to analyze features that distinguish the rotor vibration as well as stator current samples of faulty induction motor from the healthy one.

The parameters used and their levels are shown in Table 2. Totally three tests were performed in this experiment for both vibration and current samples. Test 1 for both vibration and current samples used the set $N_1 \cup F_1$ as set, in which there are totally 62 numbers of samples and out of these, 31 samples were considered as the training set and another 31 samples were used as test set. Test 2 for both vibration and current samples used the set $N_2 \cup F_2$ as set, in which there are totally 62 numbers of samples and out of these, again 31 samples were considered as the training set and another 31 samples were used as test set. Test 3 used $N_1 \cup F_1$ as the training set and the set $N_2 \cup F_2$ as the testing set. In each test, backpropagation neural network was implemented to compare the performances using multiscale entropy-related features, described earlier.

In this work, the number of inputs to the classifier is two (speed and fault) and the output is one (fault detection using either vibration or current). In case of vibration, the architecture 2–6–1 gives the minimum MSE and is considered best for all three tests, and in case of current samples, the architecture 2–13–1 is best for the test 1 and test 2 (as shown in Fig. 6), but in case of third test, the architecture 2–

15–1 is considered best. The above-said network was then trained and tested for various values. The optimal network for the broken rotor bar fault detection using vibration and current is shown in Tables 3, 4, 5, 6, 7 and 8. Best network architecture which is giving minimum error in test condition (in Tables 3, 4, 5, 6, 7 and 8) are marked bold.

Fig. 6 Variation of mean square error with number of neurons in hidden layer used for current of test 2

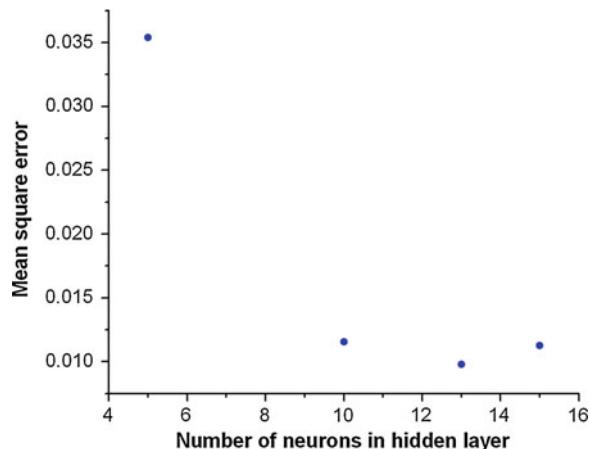


Table 3 Prediction accuracy on the test set for test 1 of vibration signal

Network architecture	Number of iterations	Mean square error training	Mean square error testing	Maximum predicted error (%)
2-5-1	1,307	0.0092	0.0152	14.28
2-6-1	1,059	0.0057	0.0100	12.16
2-9-1	5,556	0.0124	0.02349	22.751

Test 1 = $N_1 \cup F_1$ includes totally 62 samples, training set = 31 samples and testing set = another 31 samples

Table 4 Prediction accuracy on the test set for test 2 of vibration signal

Network architecture	Number of iterations	Mean square error training	Mean square error testing	Maximum predicted error (%)
2-4-1	30,331	0.0350	0.0275	19.097
2-6-1	5,101	0.01011	0.020	13.88
2-8-1	12,978	0.0190	0.022	17.41

Test 2 = $N_2 \cup F_2$ includes totally 62 samples, training set = 31 samples and testing set = another 31 samples

4.1 Fault Detection Using Vibration Signals

See Tables 3, 4, and 5; Fig. 7.

Table 5 Prediction accuracy on the test set for test 3 of vibration signal

Network architecture	Number of iterations	Mean square error training	Mean square error testing	Maximum predicted error (%)
2-4-1	12,000	0.0179	0.0196	19.844
2-6-1	1,083	0.0115	0.0116	15.511
2-8-1	1,915	0.102	0.0261	22.14

Test 3 = $N_1 \cup F_1$ as training set and $N_2 \cup F_2$ as testing set

Table 6 Prediction accuracy on the test set for test 1 of current signal

Network architecture	Number of iterations	Mean square error training	Mean square error testing	Maximum predicted error (%)
2-10-1	12,722	0.0159	0.0208	13.8732
2-13-1	696	0.0101	0.0077	11.521
2-15-1	10,722	0.0110	0.0164	13.088

Test 1 = $N_1 \cup F_1$ includes totally 62 samples, training set = 31 samples and testing set = another 31 samples

Table 7 Prediction accuracy on the test set for test 2 of current signal

Network architecture	Number of iterations	Mean square error training	Mean square error testing	Maximum predicted error (%)
2-10-1	1,394	0.0128	0.0126	13.95
2-13-1	1,002	0.01027	0.0094	10.99
2-15-1	1,575	0.0122	0.0115	14.056

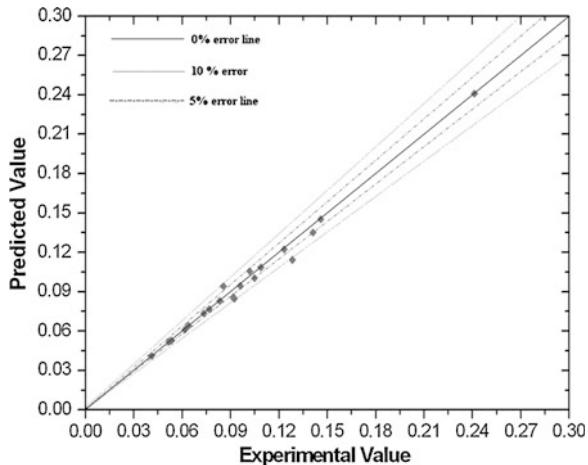
Test 2 = $N_2 \cup F_2$ includes totally 62 samples, training set = 31 samples and testing set = another 31 samples

Table 8 Prediction accuracy on the test set for test 3 of current signal

Network architecture	Number of iterations	Mean square error training	Mean square error testing	Maximum predicted error (%)
2-10-1	12,430	0.016	0.0196	19.889
2-15-1	4,400	0.010	0.01107	14.01
2-17-1	10,245	0.012	0.019983	17.099

Test 3 = $N_1 \cup F_1$ as training set and $N_2 \cup F_2$ as testing set

Fig. 7 Variation of predicted value with experimental value used for current of test 2



4.2 Fault Detection Using Current Signals

See Tables 6, 7, and 8.

The comparison between the experimental values and results obtained by BPNN for test 2 of current is shown in Fig. 6. From the above given results, it can be observed that the predicted vibration is about 15.5 % for all cases and predicted result for current is maximum within 14.01 % of the experimental value. The present study clearly shows that the BPNN model can be trained to predict the broken rotor bar fault using either vibration or current with reasonable accuracy.

5 Conclusions

This work presents the detailed information about the impact of broken rotor bar on the rotor vibration and stator current response by using statistical analysis. The aim of the corresponding experimental investigations is to contribute toward the improvement in the rotor fault diagnosis by the use of MSE.

The following conclusions are derived on the basis of investigation:

1. The experiments carried out in this paper may contribute in the diagnosis of rotor fault by employing the rotor vibration as well as stator current.
2. The experimental results and the values predicted by BPNN model are fairly close within 15.5 % for vibration and 14.01 % for current signals, which shows that the vibration and current during rotor fault detection can be effectively modeled using BPNN.
3. From the above results and discussions, it may be found that stator current alone may be used for the rotor fault detection.

References

1. Bonnet, A.H.: Analysis of rotor failures in squirrel cage induction machines. *IEEE Trans. Ind. Appl.* **24**(6), 1124–1130 (1988)
2. Bonnet, A.H., Soukup, G.C.: Cause and analysis of stator and rotor failures in three-phase squirrel-cage induction motors. *IEEE Trans. Ind. Appl.* **28**(4), 921–937 (1992)
3. Nandi, S., Toliyat, H.A.: Fault diagnosis of electrical machine—a review. In: Proceedings of International Electric Machines and Drives Conference (IEMDC), May 1999, Seattle, WA, pp. 219–221 (1999)
4. Thomson, W.T., Fenger, M.: Current signature analysis to detect induction motor faults. *IEEE Ind. Appl. Mag.* **7**(4), 26–34 (2001)
5. Maier, R.: Protection of squirrel-cage induction motor utilizing instantaneous power and phase information. *IEEE Trans. Ind. Appl.* **28**, 376–380 (1992)
6. Legowski, S.F., Ula, S., Trzynadlowski, A.M.: Instantaneous stator power as a medium for the signature analysis of induction motors. *IEEE Trans. Ind. Appl.* **32**, 904–909 (1996)
7. Hsu, J.S.: Monitoring of defects in induction motors through air-gap torque observation. *IEEE Trans. Ind. Appl.* **31**, 1016–1021 (1995)
8. Yan, R.Q., Gao, R.X.: Approximate entropy as a diagnostic tool for machine health monitoring. *Mech. Syst. Signal Process.* **21**, 824–839 (2007)
9. Yan, R.Q., Gao, R.X.: Machine health diagnosis based on approximate entropy. In: Proceedings of ICMT 2004 Instrumentation and Measurement Technology Conference, Italy, pp. 2054–2059 (1995)
10. Pincus, S.M.: Approximate entropy as a measure of system complexity. *PNAS* **88**(6), 2297–2301
11. Richman, J.S., Moorman, J.R.: Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. H.* **278**, 2039–2049 (2000)
12. Haitham, M.A., Alan, V.S.: Use of sample entropy approach to study heart rate variability in obstructive sleep apnea syndrome. *IEEE Trans. Bio. Eng.* **50**, 1900–1904 (2002)
13. Costa, M., Goldberger, A.L., Peng, C.K.: Multiscale entropy analysis of complex physiologic time series. *Phys. Res. Lett.* **89**(6), 68–102 (2002)
14. Fan, X.F., Zuo, M.J.: Machine fault feature extraction based on intrinsic mode functions. *Meas. Sci. Technol.* **19**(4), 045105 (12 p) (2008)
15. Zhang, L., Xiong, G., Liu, H., Zou, H., Guo, W.: An intelligent fault diagnosis method based on multiscale entropy and SVMs. *Lecture Notes in Computer Science*, vol. 5553, pp. 724–732 (2009)
16. Lin, J.L., Liu, J.Y.C., Li, C.W., Tsai, L.F., Chung, H.Y.: Motor shaft misalignment detection using multiscale entropy with wavelet denoising. *Expert Syst. Appl.* **37**, 7200–7204 (2010)
17. Zhang, L., Xiong, G., Liu, H., Zou, H., Guo, W.: Bearing fault diagnosis using multi-scale entropy and adaptive neuro-fuzzy inference. *Expert Syst. Appl.* **37**, 6077–6085 (2010)
18. Schoen, R.R., Habetler, T.G., Kamran, F., Bartheld, R.G.: Motor bearing damage detection stator current monitoring. *IEEE Trans. Ind. Appl.* **31**, 1274–1279 (1995)

Design of a Signal Sensor for Analyzing Biological Activities at Cellular Level

Amit Krishna Dwivedi, Anubhav Sinha and Aminul Islam

Abstract Potentiostat is a circuit arrangement that maintains the electrochemical stability and also buffers the output signal. It senses the signal applied to it and thus generates the output proportional to the electrochemical current. This paper proposes an efficient design of a low-amplitude signal sensor with very low power consumption. The proposed design proves its significance as a low-amplitude signal sensor, which finds its applications in the field of biomedical. Different types of low-amplitude signals are processed through the proposed potentiostat design to analyze the sensed output levels. Further, this paper also makes the variability analysis of proposed design to demonstrate its immunity against the process parameter variation. The modeling of proposed potentiostat is done in SPICE, and the simulation results have been extensively verified using the same.

Keywords Biomedical applications · Low-amplitude sensor · Potentiostat

1 Introduction

Sometimes, it is our requirement to measure very low-amplitude current values. Especially, in case of the biochemical reactions, there is a flow of current due to difference in the concentration of the fluids present in the human body. As the pH value of the different fluids present in our body varies, hence, there is a flow of very small amplitude of current in the body. In the biomedical applications, we require these measurements for precise and accurate analysis. Hence, this paper presents a

A.K. Dwivedi (✉) · A. Sinha · A. Islam
Birla Institute of Technology, Mesra, Ranchi 835215, Jharkhand, India
e-mail: amit10011.13@bitmesra.ac.in

A. Sinha
e-mail: anubhav10461.12@bitmesra.ac.in

A. Islam
e-mail: aminulislam@bitmesra.ac.in

low power consumption, highly precise, and compact design of a potentiometer that can be used to detect very small flow of currents.

Potentiostats are the electronic circuit arrangements designed to detect many biologically and environmentally important analytes [1]. Properties of analytes and its chemical constituent's measurements are of interest in an analytical procedure. Measurement of weak currents enables us to predict and conclude several properties of biological activities at the cellular level. The state of the art for the potentiostats is based upon the measurement of the charge flow between the different electrodes. The scaling of MOS transistors limits the performance of the potentiostats. Thus, there is demand for increased functionality and reduced system size potentiostats. Amperometry involves the chemistry of measurement of flow of ions. The drive voltages of amperometric chemical sensors are defined as the reduction/oxidation (redox) potentials of the analytes. Traditionally reported low power consumption potentiostats designed with op-amps suffer major problem of device mismatch. They need to be calibrated using offset correction [2]. These reported designs also possess large electrode capacitances. This results in instability of these potentiostats [3]. Apart from these issues, we require compact and low power consumption with immunity against the process parameter variations. The proposed design in this paper overcomes these demerits and also proves immune against the process parameter variations. This paper makes the following contributions:

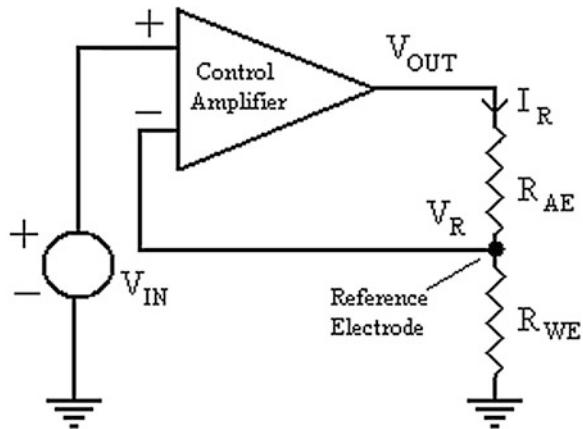
- (1) In view of the aforesaid requirement, this paper proposes a MOS-based design of potentiostat.
- (2) The variability analysis of the various design metrics of the proposed design is also presented in this paper.

Rest of this paper is organized as follows: Section 2 presents the proposed design of potentiostat. Simulation results and discussions are mentioned in Sect. 3. Comparative study of the proposed design is mentioned in Sect. 4. Finally, concluding remarks are made in Sect. 5.

2 Proposed Design

A simple arrangement of amperometric sensor is designed to detect the flow of ions by means of electric currents [4]. It typically consists of three types of electrodes—auxiliary electrode (AE), working electrode (WE), and reference electrode (RE). The potentiostat is an electronic system used to control the three-electrode system. The three-electrode system maintains a constant potential between WE and RE by controlling current at AE [5]. It also ensures that no current is drawn at the RE and all reduction and oxidation reactions occur at the WE. Figure 1 shows the basic structure of potentiostat with control amplifier and different electrode's resistances. The control amplifier traces the bias input signal (V_{IN}) by maintaining the voltages at RE and WE which is done by controlling current (I_R) at the output terminal. Relation between V_R and V_{IN} can be expressed using following equations:

Fig. 1 Basic structure of potentiostat with control amplifier and different electrode's resistances



$$V_{\text{OUT}} = A(V_{\text{IN}} - V_R) \quad (1)$$

$$I_R = \frac{V_{\text{OUT}}}{R_{\text{AE}} + R_{\text{WE}}} = \frac{V_R}{R_{\text{WE}}} \quad (2)$$

$$V_R = \frac{R_{\text{WE}}}{R_{\text{WE}} + R_{\text{AE}}} V_{\text{OUT}} \quad (3)$$

where A represents the gain of the control amplifier, R_{WE} is the resistance of the WE, R_{AE} is resistance of the AE, and I_R is the electrochemical current. Here, I_R is proportional to the analyte's concentration and comes into picture as soon as cell voltage (V_{cell}) approaches an analyte's redox potential. Equation (2) evaluates the current I_R with assumption that no current is drawn by RE. Now, from (1) and (3), we can write as follows:

$$V_R = \frac{A \left(\frac{R_{\text{WE}}}{R_{\text{WE}} + R_{\text{AE}}} \right)}{1 + A \left(\frac{R_{\text{WE}}}{R_{\text{WE}} + R_{\text{AE}}} \right)} V_{\text{IN}} \Rightarrow V_R = V_{\text{IN}} \quad (4)$$

As value of $A(R_{\text{WE}}/R_{\text{WE}} + R_{\text{AE}}) \gg 1$, hence, Eq. (4) shows that V_R traces V_{IN} by controlling current I_R . Measurement of current I_R gives the value of current flowing through the electrodes due to developed electrode's resistance.

$$V_{\text{cell}} \equiv I_R \times R_{\text{WE}} \quad (5)$$

Thus, an electrical circuit equivalent to an amperometric sensor can be represented as a simplified two-resistor model which is utilized in this paper. A fully differential design of a potentiostat circuit consists of a differential input amplifier which is used to compare the potential between the RE and the WE [4, 5]. The RE is used to indicate the bias potential. This paper proposes an alternate design of the

bioelectronic potentiostat system which can be utilized to sense very low amplitude of currents. Figure 2 shows the MOS-implemented circuit-level model of the proposed potentiostat design. In the proposed design, two inverters (inverter with transistors P1 and N1 and inverter with transistors P2 and N2) are cascaded to control the flow of current through the auxiliary node. N-times cascaded inverter sets can be incorporated in the proposed design to increase the accuracy in terms of tracing the exact input signal. We have utilized two sets of cascaded inverters to develop the results presented in this paper. The differential input terminals (+ve signal at MN2 and -ve signal at MP2) correspond to the WE. The output voltage

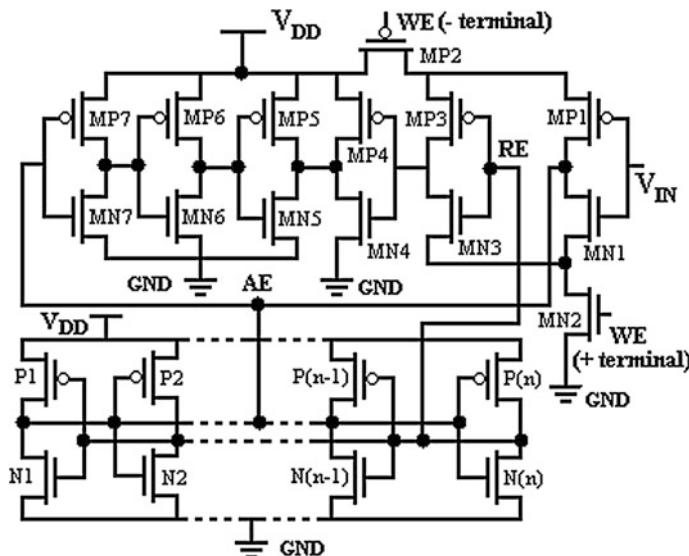
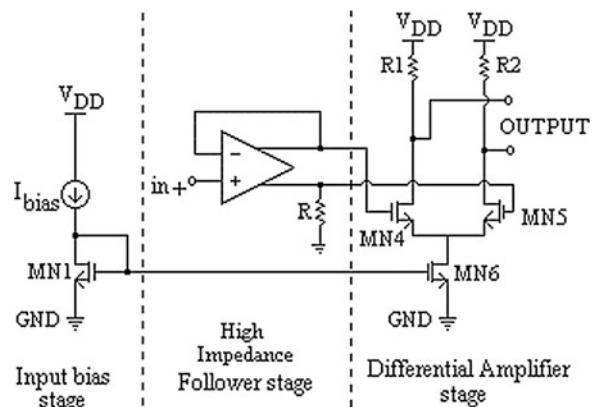


Fig. 2 MOS-implemented circuit-level model of the proposed potentiostat design

Fig. 3 Amplification scheme for the proposed circuit-level model of the potentiostat



swing of the proposed design is increased significantly by incorporating an additional amplifier circuitry. Figure 3 shows the amplification scheme for the proposed design. The proposed amplification scheme is designed to incorporate 40 dB of gain to the input sensed signals. The design consists of basically three stages of amplification—(i) input bias circuitry, (ii) high-impedance follower stage, and (iii) differential amplifier stage. The differential amplifier scheme employed here helps in common mode rejection [6]. It also helps in achieving amplified output waveforms, proportional to the applied signal to its WE terminals.

3 Simulation Results and Discussion

An efficient design of a potentiostat capable of sensing very low-amplitude signal is presented. This paper demonstrates the electrochemical analyses using the proposed design of potentiostat by sensing signals resembling with the electrochemistry. Figure 4 shows the output waveform obtained after the application of square waveform as an input signal to the proposed potentiostat. The Fig. 4a shows the waveform obtained at the output terminal, while Fig. 4b is showing the traces of the input signal obtained at the RE (see Fig. 4) which is maintained by controlling the current through the AE. Similarly, Fig. 5 shows a triangular waveform being sensed by the proposed potentiostat design and the input signal being traced at the RE.

The proposed potentiostat design has been simulated using predictive technology model (PTM) of MOSFET@ 16-nm technology node [7]. This paper also makes the variability analysis of the different design metrics for the proposed potentiostat design. Table 1 shows the numeric data of these metrics for two different input signals.

The variability analysis of the response time (delay in response) and power consumption of the proposed potentiostat for different input signals are mentioned. The analysis is done by varying device parameters such as channel length (L), gate width (W), channel doping concentration (N_{DEP}), oxide thickness (t_{OX}), threshold

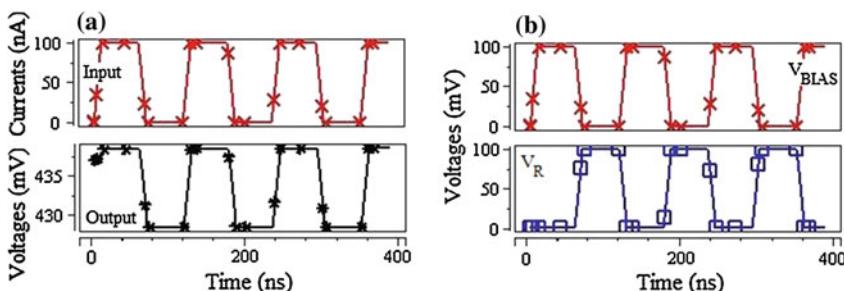


Fig. 4 **a** Output waveform obtained from the proposed potentiostat design after application of square pulses. **b** V_{IN} is traced at the RE after some delay

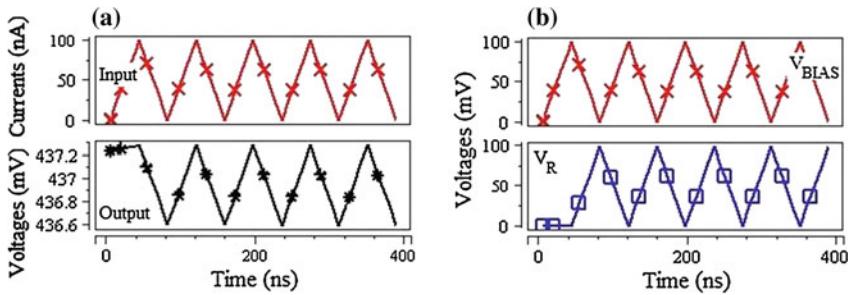


Fig. 5 **a** Output waveform obtained from the proposed potentiostat design after application of triangular pulses. **b** V_{IN} is traced at the RE after some delay

Table 1 Analysis of the different parameters of the proposed potentiostat design for input signals

V_{DD} (volts)	Input triangular signal			Input trapezium signal		
	Delay (s)	Power (W)	PDP (W-s)	Delay (s)	Power 2(W)	PDP (W-s)
0.65	3.8e-10	4.8e-06	1.8e-15	5.0e-10	4.6e-06	2.3e-15
0.675	4.0e-10	6.2e-06	2.5e-15	5.2e-10	6.0e-06	3.1e-15
0.70	4.1e-10	8.0e-06	3.3e-15	5.3e-10	7.7e-06	4.1e-15
0.725	4.4e-10	1.0e-05	4.4e-15	5.4e-10	9.9e-06	5.3e-15
0.750	4.6e-10	1.3e-05	5.9e-15	5.6e-10	1.3e-05	7.3e-15

voltage (V_t), carrier mobility (μ_0), and supply voltage (V_{DD}) up to $\pm 10\%$. For higher accuracy, 5,000 samples of the proposed circuit-level model have been simulated using Monte Carlo simulations.

4 Comparative Study of the Proposed Design

The proposed fully differential design of the potentiostat is capable of controlling both AE's and WE's voltages to achieve twice the voltage range as compared to its single-ended design [3]. The differential inputs design also reduces the susceptibility to common mode interference [6]. The proposed circuit employs high input impedance state to make measurements accurate, and it also ensures that no current is drawn on the RE. N -times cascading the two sets of cascaded inverter help controlling the AE. The CMRR values of the potentiostat design reported in [8] are limited by the resistor mismatching as it is difficult to exactly match the analog components. The output impedance and the transconductance gain of the CMOS-based potentiostat design reported in [9] depend on the single transistor which is prone for mismatch due to bias point instability. The reported design also shows the trade-off between r_{o2} and g_{m2} which limits the gain of the proposed design.

Table 2 Design specification of the proposed potentiostat

Specification	Numeric values
Range of sense current	0.1 pA–100 nA
Input bias voltage (V_{IN})	0.7–1.0 V
Amplification of the output signal	40 dB
Maximum output swing	2 × Input bias voltage (V_{IN})

However, these above-mentioned issues are avoided in the proposed design by utilizing a network of transistors in push-pull arrangements which avoids the mismatch errors. Further, amplification scheme proposed in this paper is capable of boosting the sensed signal with gain of 40 dB. Different design metrics for the proposed potentiostat have been tabulated in Table 2.

5 Conclusion

A fully differential input, high-gain, and low-input impedance potentiostat design is presented. The proposed design is capable of detecting low-amplitude signals such as current flow in electrochemical reactions or detecting biological activities at the cellular level or amperometric analysis at the electrodes. The differential input design minimizes the common mode interferences and also reduces mismatch errors. Different input signals of very low amplitude in the order of micro-or nano-levels are detected by utilizing the proposed design. The amplification scheme for the same is also presented in this paper. Thus, with further optimization, exclusive utilization of the proposed potentiostat is presented. The proposed design also proves efficient in terms of immunity against the process parameter variations which is demonstrated by variability analysis.

References

1. Kissinger, P., Preddy, C., et al.: Overcoming Solution Resistances with Stability and Grace in Potentiostatic Circuits. In: Kissinger, P., Heineman, W. (eds.) *Laboratory Techniques in Electroanalytical Chemistry*, Revised and Expanded, 2nd edn. Marcel Dekker, CRC Press, New York (1996)
2. Duwe, M., Chen, T.: Low power integrated potentiostat design for μ electrodes with improved accuracy. In: IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS), pp. 1–4 (2011)
3. Fidler, J.C., Penrose, W.R., Bobis, J.P.: A potentiostat based on a voltage-controlled current source for use with amperometric gas sensors. *IEEE Trans. Instrumntation Measur.* **41**(2), 308–310 (1992)
4. Grieshaber, D., MacKenzie, R., Voros, J., Reimhult, E.: Electrochemical biosensors—sensor principles and architectures. *Sensors* **8**, 1400–1458 (2008)

5. Lim, S., Choi, I., Lee, H: Biochemical sensor interface circuits with differential difference amplifier. In: IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), pp. 176–179 (2012)
6. Ramirez-Angulo, J., Balasubramanian, S., Lopez-Martin, A.J., Carvajal, R.G.: Low voltage differential input stage with improved CMRR and true rail-to-rail common mode input range. *IEEE Trans. Circuits Syst. II Expr. Briefs* **55**, 1229–1233 (2008)
7. Berkeley Predictive Technology Model, UC Berkeley Device Group. [Online]. Available <http://www-device.eecs.berkeley.edu/~ptm/>
8. Martin, S.M., Gebara, F.H., Strong, T.D., Brown, R.B.: A low-voltage, chemical sensor interface for systems-on-chip: the fully-differential potentiostat. In: IEEE Proceedings of the International Symposium on Circuits and Systems (ISCAS'04), pp. 892–895, 23–28 May 2004
9. Ayers, S., Gillis, K.D., Lindau, M., Minch, B.A.: Design of a CMOS potentiostat circuit for electrochemical detector arrays. *IEEE Trans. Circ. Syst. I Regul. Pap.* **54**, 736–744 (2007)

Fuzzy PID Control for Ball and Beam Mechanism

Nirmal Baran Hui and Pratima Sharma

Abstract A ball beam mechanism is found to be highly unstable system. PID control has been mostly used for stabilizing the same. However, tuning of PID controller gains is a difficult issue. In this work, attempt has been made to develop a fuzzy model-based ball and beam mechanism. Initially, an analytical model of the system is presented in the nonlinear state space form. The feasibility of the proposed scheme for ball and beam system is demonstrated using MATLAB/SIMULINK. Computer simulations are exploited to demonstrate the validity and feasibility of the developed control schemes.

Keywords Ball beam model · PID control · Fuzzy logic controller

1 Introduction

The ball and beam control system is an example for a highly nonlinear problem with one degree of freedom. When such a system is simulated, the ball falls quickly and this makes the control very difficult. From the literature, it is seen that researchers were given more importance in designing of controllers and PID tuning with different techniques [1–7]. For the past few years, intelligent computational techniques were utilized effectively to solve control design and modeling problems of ball and beam system. Among them, fuzzy systems, neural networks, and genetic algorithms are the most used methods. Some investigations are reported to have addressed the modeling and control of a ball and beam system using various model-based and artificial intelligence (AI)-based approaches.

N.B. Hui (✉) · P. Sharma

Department of Mechanical Engineering, National Institute of Technology Durgapur,
Durgapur 713209, West Bengal, India
e-mail: nirmalhui@gmail.com

P. Sharma
e-mail: pratru171@gmail.com

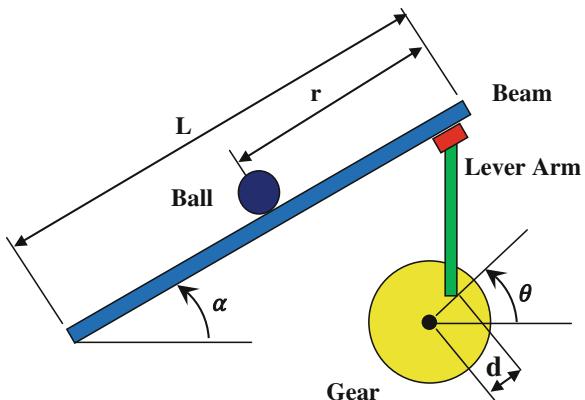
The idea of fuzzy modeling was first proposed by Wenbiao [8] and has subsequently been pursued by many others for more than two decades. The fuzzy modeling for complex processes is regarded as one of the key problems in fuzzy systems research [8, 9]. In the field of fuzzy modeling, the Takagi-Sugeno (T-S) fuzzy model [10, 11] has been used to approximate accurately the dynamics of complex systems. Besides the capability of modeling nonlinear systems, there are other properties that make fuzzy models interesting not only theoretically but also for the industrial practice.

2 Mathematical Models and Control of a Ball and Beam Mechanism

The ball and beam system that is shown in Fig. 1 consists of a long beam which can be tilted by a servo or electric motor together with a ball rolling back and forth on top of the beam. The significance of the ball and beam system is that it is a simple system which is open-loop unstable. Even if the beam is restricted to be nearly horizontal without active feedback, it will swing to one side or the other, and the ball will roll off to the end of the beam. To stabilize the ball, a control system which measures the position of the ball and adjusts the beam accordingly is used. In two dimensions, the ball and beam system becomes the ball and plate system, where a ball rolls on top of a plate whose inclination can be adjusted by tilting it frontward, backward, leftward, or rightward.

For this problem, we will assume that the ball rolls without slipping, and friction between the beam and ball is negligible. Mass, radius, moment of inertia, and position of the ball are denoted by M , R , J , and r , respectively. Length and angle coordinates of the beam are expressed using L and α , respectively. Lever arm offset is denoted by d , and θ represents the servo gear angle.

Fig. 1 Ball beam mechanism



2.1 System Equations

The Lagrangian equation of motion for the ball is given by the following:

$$\theta = \left(\frac{J}{R^2} + M \right) \ddot{r} + Mg \sin \alpha - Mr(\dot{\alpha})^2 \quad (1)$$

Linearization of this equation about the beam angle, $\alpha = 0$, gives us the following system:

$$\left(\frac{J}{R^2} + M \right) \ddot{r} = -Mg\alpha \quad (2)$$

The equation which relates the beam angle to the angle of the gear can be approximated as linear by the equation below:

$$\alpha = \frac{d}{L} \theta \quad (3)$$

Substituting this into the previous equation, we get:

$$\left(\frac{J}{R^2} + M \right) \ddot{r} = -\frac{Mgd}{L} \theta \quad \text{or} \quad \frac{r}{\theta} = \frac{-\frac{Mgd}{L}}{\left(\frac{J}{R^2} + M \right) s^2} \quad (4)$$

2.2 SIMULINK Model

Initially, a simulation model was constructed using MATLAB/SIMULINK as shown in Fig. 2. It has been observed that the system represented in Eq. (4) is an unstable system (refer to Fig. 3) with the following system parameters: $M = 0.11 \text{ kg}$; $R = 0.015 \text{ m}$; $d = 0.03 \text{ m}$; $g = 9.8 \text{ m/s}^2$; $L = 1.0 \text{ m}$; $J = 9.99e-6 \text{ kg m}^2$. In order to stabilize and control the said system, classical PID control strategy has been applied. The PID-controlled ball and beam system is shown in Fig. 4. Performance

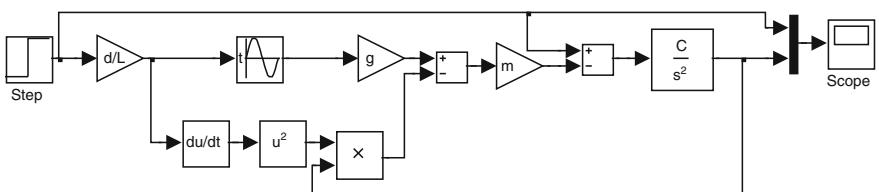


Fig. 2 SIMULINK model of a ball and beam system

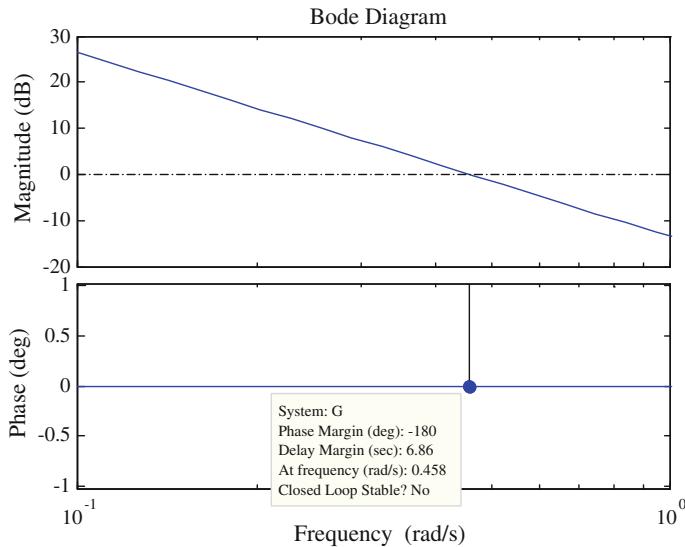


Fig. 3 Bode diagram corresponding to the transfer function represented by Eq. (4)

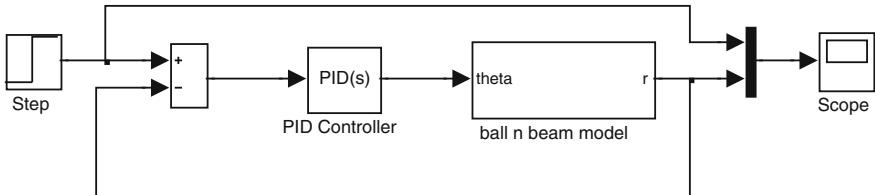


Fig. 4 SIMULINK model of the PID-controlled ball and beam system

of PID control law largely depends on its gain values, and Ziegler–Nichols method-II has been applied to get controller gain values.

2.3 Fuzzy Logic-Based Tuning of PID Controller

Tuning of PID controller gains using Ziegler–Nichols method does not provide satisfactory result. It may be due to following reasons: (a) Constant gain values throughout the simulation time do not provide good result due to the nonlinearity present in the model; (b) It lacks robustness and adaptability.

Keeping in view of the above-listed drawbacks of a PID controller, whose gains have been tuned using Ziegler–Nichols method, an attempt has been made to tune the controller gains using a fuzzy logic controller (FLC). The FLC works on fuzzy

set theory proposed by Wenbiao. There are two major approaches of developing FLC: (i) Mamdani approach [10] (ii) T-S approach [11]. In the present study, the PID controller was tuned using an FLC following Mamdani approach and simulation was carried out using MATLAB. The two condition variables used are: *error* (ϵ) in pendulum angle (difference between desired angular position of pendulum and the present position) and *differential change in error* (derivative of error, $\dot{\epsilon}$). The output consists of three gain values K_P , K_D , and K_I . The shape of membership functions is assumed to be triangular for simplicity. Figure 5 shows the author-defined membership function distribution of both input as well as output variables. Five linguistic terms Zero (Z), Low (L), Medium (M), High (H), and Very High (VH) have been considered for all the outputs and five linguistic terms have also been considered for outputs such as Negative High (NH), Negative Low (NL), Zero (Z), Positive Low (PL), and Positive High (PH). With five choices each for *error* and *d_error*, there could be 25 input combinations. For each of these 25 input

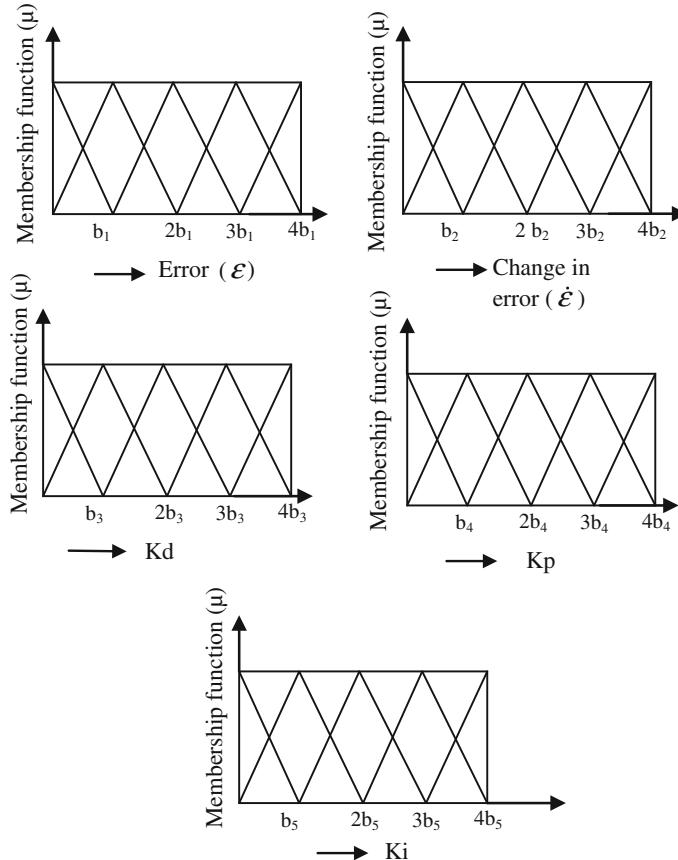


Fig. 5 Membership function distribution of FLC

Table 1 Rule base of FLC

No.	Input combinations		Output		
	ε	$\dot{\varepsilon}$	K_p	K_i	K_d
1	NH	NH	Z	H	L
2	NH	NL	Z	H	Z
3	NH	Z	M	M	M
4	NH	PL	L	L	M
5	NH	PH	M	Z	L
6	NL	NH	L	L	Z
7	NL	NL	Z	L	Z
8	NL	Z	L	L	H
9	NL	PL	L	Z	H
10	NL	PH	Z	L	M
11	Z	NH	VH	Z	VH
12	Z	NL	L	H	L
13	Z	Z	L	M	M
14	Z	PL	H	L	M
15	Z	PH	Z	M	M
16	PL	NH	L	M	M
17	PL	NL	L	L	L
18	PL	Z	Z	M	H
19	PL	PL	L	L	M
20	PL	PH	H	L	Z
21	PH	NH	H	H	M
22	PH	NL	H	Z	M
23	PH	Z	L	M	M
24	PH	PL	Z	H	L
25	PH	PB	L	L	VH

combinations, there could be three gain values, one each for K_p , K_D , and K_I . Thus, there could be at least 25 rules in the rule base. These rules are set manually based on intuition and are presented in Table 1.

One such rule could be as follows:

IF $error$ is **NL** and d_error is **Z**, THEN K_p is **L**, K_I is **L** and K_D is **H**.

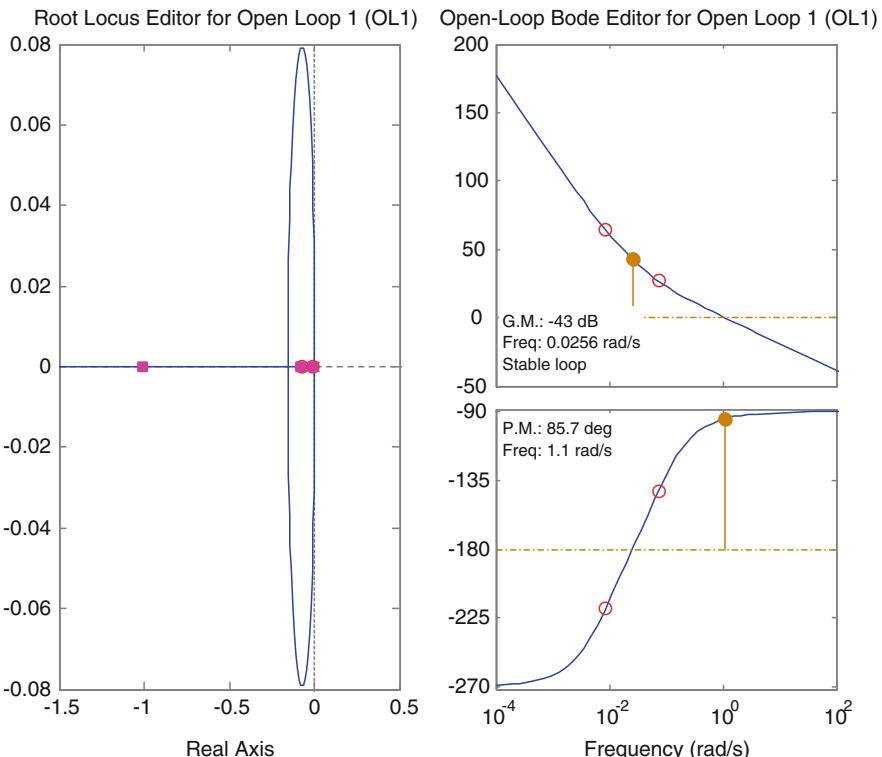
3 Results and Discussions

In the present study, the controlling action has been incorporated to the system using SISO design tool box in MATLAB environment. Classical PID control technique has been adopted in this case. The SISO tool provides the means for

Table 2 Gain margin (GM) and phase margin (PM) of the system for different controllers

Controller type	Ziegler–Nichols-based PID tuning (frequency response-based)		Robust response time method	
	GM (dB)	PM (deg)	GM (dB)	PM (deg)
PID	3.41 @ 0.00204 rad/s	11.8 @ 0.0025 rad/s	34.2 @ 0.0985 rad/s	60 @ 1 rad/s
PID with derivative filter	4.45 @ 0.00194 rad/s	16 @ 0.00256 rad/s	43 @ 0.0256 rad/s	85.7 @ 1.1 rad/s

automated PID tuning, and it automatically generates the compensator value. Table 2 shows the stability margin values for different controllers tuned using two methods, out of which robust response time (RRT) method provided the best stability margin (refer to Table 2; Fig. 6). Finally, K_P , K_D , and K_I values are extracted corresponding to the best compensator, which are as follows: $K_P = -8,675.56$, $K_D = -100,862.2$, and $K_I = -70.53$. It is also to mention that the negative values of the compensator gains indicate the use of direct acting controller.

**Fig. 6** Root locus and Bode plot for the system using PID controller tuned by RRT method

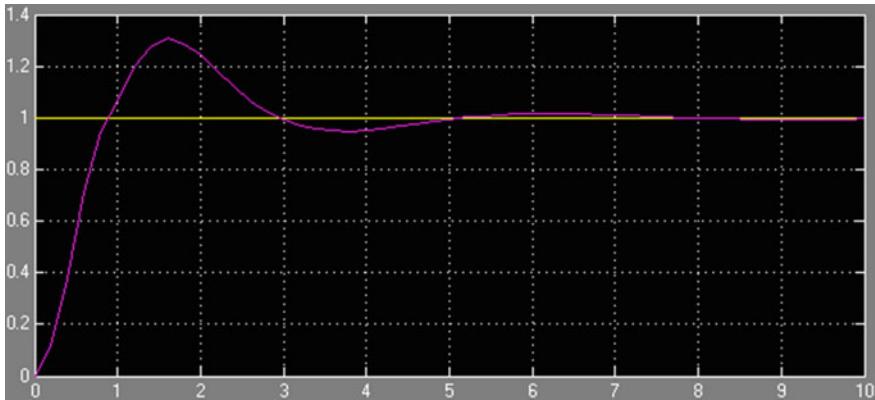


Fig. 7 Ball position under step input using PID controller tuned by RRT method

It indicates an increase in the controlled variable requires an increase in the control action. Otherwise, the feedback loop should be made positive.

Finally, comparative analysis has been made between the two approaches. In the first approach, PID controller gains tuned using RRT method, and in the next approach, those were tuned by an Fuzzy Logic (FL). Figures 7 and 8 show the position of the ball under two different approaches, respectively. It has been found that second approach has taken less time to settle down (3 s compared to the 4.8 s of the first approach). Also, percentage overshoot for the first approach is larger than the second approach. It may be due to the fact that the first approach is more adaptive in nature.

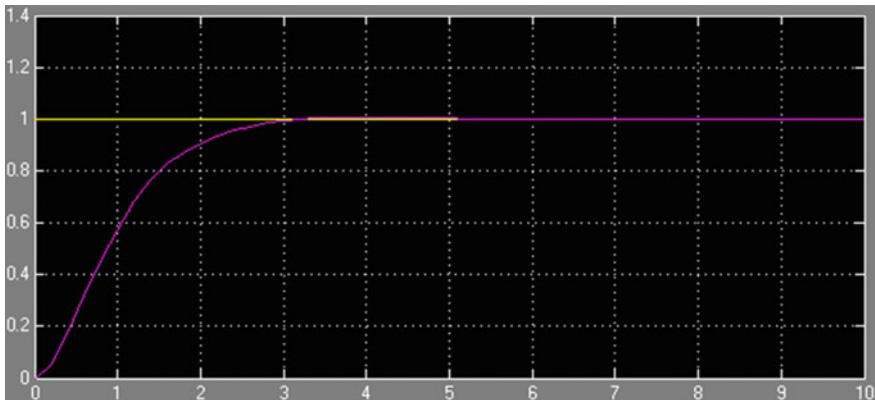


Fig. 8 Ball position under step input using PID controller tuned by FL method

4 Conclusions

An attempt has been made to test the performance of FLC on a highly instable system. In normal condition ball beam mechanism that has been considered in this paper is not stable. Main aim of this work was to stop the movement of the ball at any distinct location. Traditional PID control laws have been implemented initially, and it has been seen that it is taking more time to settle. However, increase in settling time might cause faulty behavior of the system and will be dangerous with the presence of some external disturbance. Moreover, in traditional PID control laws, controller gains are constants which do not provide any adaptiveness. Understanding all those drawbacks, FLC has been implemented to tune the controller gains. Here with the time, gain values will change. The high amount of improvement has been observed with the implementation of FLC. However, performance of FLC largely depends on its knowledge base. Optimization of which is required to get better result. Authors are presently working toward those issues.

References

1. Visioli, A.: Tuning of PID controllers with fuzzy logic. *IEE Proc. Control Theor. Appl.* **148**(1), 1–8 (2001)
2. Visioli, A.: Fuzzy logic based set-point weight tuning of PID controllers. *IEEE Trans. SMC, Part-A* **29**(6), 587–592 (1999)
3. Mann, G.K.I., Hu, B.G., Gosine, R.G.: Two-level tuning of fuzzy PID controllers. *IEEE Trans. SMC, Part-B* **31**(2), 263269 (2001)
4. Jianqiang, Y., Yubazaki, N., Hirota, K.: Stabilization control of ball and beam systems. In: Proceedings of 20th NAFIPS International Conference, vol. 4, pp. 2229–2234 (2001)
5. Joo, M.G., Lee, J.S.: Hierarchical fuzzy control scheme using structured Takagi-Sugeno type fuzzy inference. In: IEEE Proceedings of Fuzzy Systems, pp. 78–83 (1999)
6. Song, F., Smith, S.M.: Applying incremental best estimate directed search to optimize fuzzy logic controllers for a ball-and-beam system. In: IEEE Proceedings of Fuzzy Systems, vol. 2, pp. 1132–1137 (2002)
7. Sugeno, M., Yasukawa, T.: A fuzzy-logic-based approach to qualitative modeling. *IEEE Trans. Fuzzy Syst.* **1**, 7–31 (1993)
8. Wenbiao, Z., Zengqi, S.: Fuzzy modeling for complex process. In: IEEE Proceedings of the 3rd World Congress on Intelligent Control and Automation, pp. 1271–2175 (2000)
9. Tanaka, K., Wang, H.O.: *Fuzzy Control Systems Design and Analysis: A Linear Matrix Inequality Approach*. Wiley, New York (2001)
10. Takagi, T., Sugeno, M.: Fuzzy identification of systems and its applications to modeling and control. *IEEE SMC* **15**(1), 116–132 (1985)
11. Selvaganesan, N., Saraswathy, R.R.: A simple fuzzy modeling of permanent magnet synchronous generator. *J. Electr. Elektrika* **11**(1), 38–43 (2009)

Framework for Smart Health: Toward Connected Data from Big Data

P. Vignesh Raja, E. Sivasankar and R. Pitchiah

Abstract Health informatics has been witnessing a tremendous modernization by leveraging the information technology and networking. Big Data tools offer a platform for organizing huge volume of data generated out of the medical informatics systems. They offer mechanism to store data in a distributed manner and offer parallel processing environment to process the large amount of data. Even though such platforms offer scalable way of managing large volume of data, those tools do not provide mechanism to get value from the large volume of data. Healthcare data is peculiar in nature because it contains many links to within themselves, such as symptoms, practitioners, and medication. Processing such data using traditional RDBMS, Big Data tools to get the hidden value from it, is cumbersome. In this paper, we propose a framework based on graph database to connect various elements of healthcare data to get more value/insight from the healthcare data.

Keywords Health informatics · Big data · Graph database · Smart health · Analytics · Smart data

P. Vignesh Raja (✉)
C-DAC, Chennai, India
e-mail: vigneshrp@cdac.in

E. Sivasankar
Department of Computer Science and Engineering, National Institute of Technology,
Trichy, India
e-mail: sivasankar@nitt.edu

R. Pitchiah
DeitY, New Delhi, India
e-mail: pitchiah@deity.gov.in

1 Preliminaries

1.1 *IT and Health Care*

Health care primarily focuses on diagnosis, treatment to ailments, and prevention of diseases by providing proper medication. Health care has reached next level of it supremacy, when information technology is applied on it to result in health informatics. Health informatics or health information systems are the modern branches that are emerged out of the fusion of computer science, information technology, social science, and health care. Advent of electronic records, use of sensors is shaping the healthcare industry into a potential place where Big data can be rightly applied. A lot of issues like efficient data management to effective decision making out of the data are still a big challenge. The advent of health informatics has been creating a tremendous change in the field of health care [1]. It started with modernizing the healthcare systems with the introduction of electronic medical records (EMR). EMR brought much digitization in the field of health care which resulted in great improvements at the hospitals, patient care methodologies, and medication [2]. Since the recent past, sensors play a vital role in the field of health care. Body sensor network (BSN) is an amalgamation of wireless wearable smart devices. Such devices are meant to be worn by the chronic patients, and the wireless smart devices transmit the symptoms to the remote location based on the configurations [3]. Such technology aims to minimize the presence of individual human caretakers for the patients. Such advancements eventually resulted in collection of plaintext data to bulky images in large volume. Apart from these technological advancements, on the other side, various standards were formulated to govern the health informatics system. Since most of the health informatics system operates in silos, they forbid the exchange of useful medical practices and medication. Non-adherence to medication is the biggest challenge in the field of health care that draws huge spending of money [4]. HL7 is a set of international standards to leverage the exchange of administrative and clinical data among the health informatics systems [5]. HIPAA oversees the privacy issues related to healthcare data. HIPAA ensures protection to health information pertaining to the individuals and designates a set of access control mechanisms on the health information. Also, such stringent access control is managed to permit the access to health information when it is needed for patient care activities and other related legal purposes [6]. Fitting mechanism that stores, transforms, processes and infers knowledge from large amount of data would alleviate such wasteful spending of money.

In the subsequent sections, observations of empirical analysis conducted on the datasets obtained from the traditional Web-based health informatics application are presented. RDBMS-based system's behavior in comparison with the Hadoop-based system is elaborated in the next section. Finally, concept of graph-based databases and its application in health informatics is presented.

1.2 RDBMS-driven Solutions

Almost all of the health informatics systems are driven by the RDBMS technologies. Deployed at the healthcare institutes, such systems tend to record volume of healthcare data. Such data stored in the distributed data store are made available to the doctors and to the paramedics. Such data range from text (health records, reports) to image (X-ray, ECG images). With such system, eventually, there is going to be a variety and volume of data. Relationship among the tables is expressed using linking primary keys of different tables. Fetching-related records from different tables are carried out through making joins among the tables. This operation is basically a processor and memory intensive [7]. Generally, RDBMS lacks fault tolerance, linear scalability, and processing of unstructured data. Also, it cannot effectively handle high concurrent reading and writing of database [8]. Also, a highly decentralized/data sharing approach is needed to exchange the valuable medication information. Current health informatics systems forbid exchange of data or when there is need for “healthcare-related knowledge” within from the existing data. Also, RDBMS systems pose a limitation of data storage as it naturally cannot handle very large amount of data. Empirical analysis conducted at the sample RDBMS application with synthetic data, subsequent findings highlight that such systems literally failed as the amount of data to be processed is so high.

1.3 Map Reduce-based Solution

Apache Hadoop is one of the open-source frameworks for distributed data storage and parallel processing of the data. It is modeled after Google’s innovations on distributed data processing, which resulted in GFS and Map Reduce [9]. Hadoop is basically a master–slave architecture, where a node called as the Name Node functions as the master. A set of slave nodes, called as Data Nodes, contains the actual data. Meta data about the data stored in distributed manner across the data nodes are kept in the Name Node. Data Nodes are naturally having better hardware configurations than the Name Node, because actual data processing is done at the data nodes. Input data that are of huge volume are split into blocks, and these blocks are stored in the data nodes such that each block with a replication factor. Map Reduce is the programming that leverages the voluminous of data available on the data nodes to be processed in parallel manner [10].

System based on Hadoop architecture performs so well when it is compared with traditional RDBMS systems. Hadoop-based architecture paves way implicitly to manage the large amount of data. With the RDBMS-based solutions, the complexity of managing the dataset is $O(n)$, whereas Hadoop-based systems take $O(n)/m$, because of task parallelism. Not only offering a mechanism to manage the large volume of data, Hadoop has provisions for scalability. Besides this, RDBMS-based solutions literally failed as the volume of data increases.

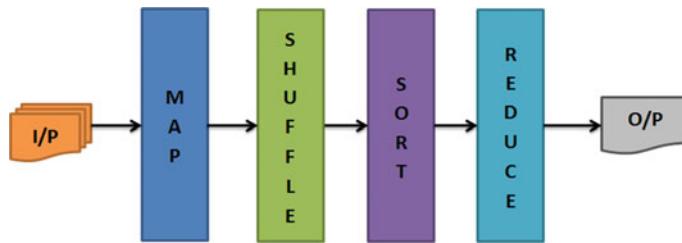


Fig. 1 Phases in Map Reduce

Obviously, Hadoop-based system has an edge over the RDBMS-based system. In spite of this, the former approach does not provide *off the shelf* mechanism to do smart analytics on the given dataset. As far as Hadoop is concerned, Map Reduce is the core to all of the data processing. Naturally, Map Reduce is the best option for processing data in a batch mode. But Map Reduce is not a choice when it comes to process the large volume of data in an iterative mode. Also, Map Reduce is not to be considered for real-time data processing. Figure 1 depicts the different phases involved in Map Reduce.

Input data are split into blocks of size 128 MB. When healthcare data of size more than TB are fed for processing, overhead is caused in the successive phases. For example, shuffle phase generates larger map outputs like many more GBs per node. When it comes to process multiple TB of healthcare data, additional customization like increasing the block size is unavoidable. Also, data transferred frequently between shuffle and sort phases add to the overhead of overall performance. When healthcare data with a list of patients along with their diseases are fed at Map Reduce system, each patient will be mapped a disease at the end of the Map phase. There is higher likelihood for the existence of duplicate entries at the end of Map phase. Let $\langle K, V \rangle$ be the set of \langle disease, patient \rangle pairs, and for the given healthcare data, there will be a list of such pairs with same disease as key but with different patients as values. There will be a list of redundant entries in the keys' list. Yet there is no provision to group all keys with same disease. Existing Map Reduce-based system does not provide mechanism to derive such links from within the healthcare data. Healthcare data are idiosyncratic in nature, because it contains multiple links within itself. Map Reduce process maps each disease from the set $\{d_1, d_2, \dots, d_n\}$ to a patient from the set $\{\text{patient}_1, \text{patient}_2, \dots, \text{patient}_n\}$ (Fig. 2).

1.4 Graph Databases

There is another innovative form of databases that concentrate on the links among the entities. Such a form of database is known as the Graph Databases. Graph databases gain relatively higher relevance in the field of health care. Graph databases contrast itself from the traditional RDBMS such that it stores not only the entity information but also relations among the entities and various properties of the

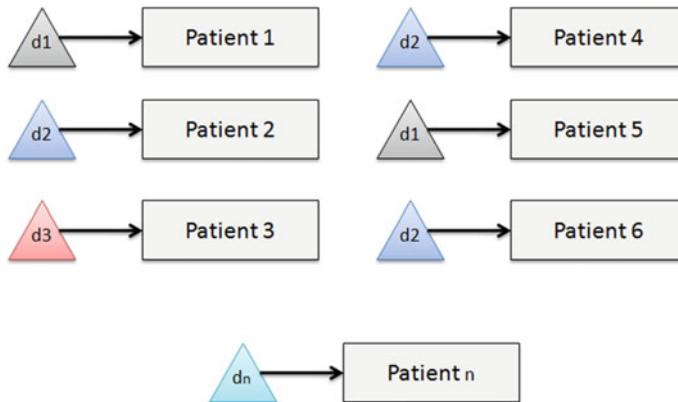


Fig. 2 Map Reduce view of healthcare data

entities. Typically, healthcare data comprises of patients, diseases, physicians, hospitals, and insurance institutes. There exists a lot of links among such data. These links can be better utilized to make much better inferences out of voluminous of healthcare data, if Graph databases are going to be employed. Graph database differs from RDBMS in the following aspects:

- Graph databases' primary objective is to manage and traverse the connected data, but RDBMS concentrates on storing and retrieving structured data.
- RDBMS' primary focus is on entities, but Graph databases intensively focus on entities, relationships among them, and its properties.
- RDBMS incurs in intensive computation to maintain relationships among entities, but Graph databases naturally maintain such relationships (Fig. 3).

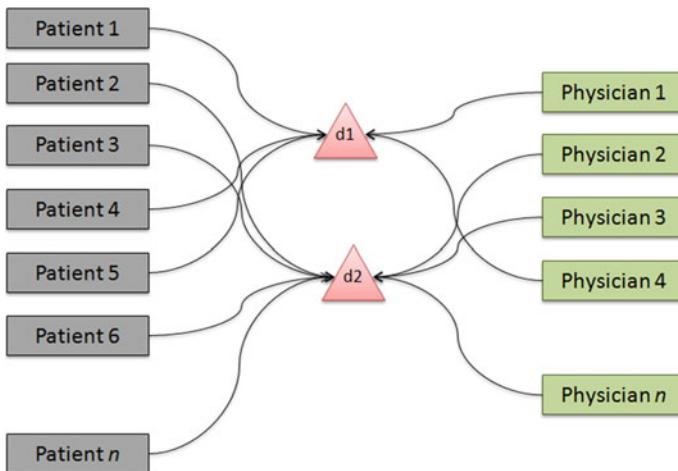


Fig. 3 Graph database view of healthcare data as connected data

Also, NoSQL databases cannot be the appropriate mechanism to express connectedness among the healthcare data. Further, requiring such facilities out of NoSQL-based solutions will add to development overheads in the application [11]. Mostly, all the healthcare data are already digitized and leveraging access to the right information at the right time is the challenge to be addressed in health informatics. Such an attempt would alleviate the increasing healthcare costs and improve the overall efficiency in health care. The essential aim out of these efforts is to achieve patient-centric smart health care. Applying Graph databases and its subsequent mining can be viewed as an enabler toward the aforementioned goal.

2 Proposed Framework

Healthcare data exist in various formats, and there are already many solutions to manage the healthcare institute operations. Most of these applications manage data on the RDBMS. Proposed framework aims to handle all the legacy data. In order to accommodate large volume of data subsequently to ensure scalability thereon, proposed framework is based on HDFS and it follows a clustered architecture. All the nodes in the resultant cluster are managed by the Hadoop framework by imposing master-slave architecture on the nodes. Such a framework aims at bestowing resources to balance large volume of data storage and computation-intensive processes. Also, the proposed framework is entirely based on open-source toolset; hence, extension of features and custom feature development is at ease.

Even though recent advancements in the field of RDBMS equipped them to support large amount of datasets, fundamentally, they are bound to specific data formats. For instance, Postgresql 9.x can handle larger volume of data (in terms of TBs), and it mandates higher hardware configuration to be available. Also, RDBMS is subject to mid-query failures, whereas Hadoop's design strategy has enabled the proposed framework to run on commodity hardware. Further, it offers better fault tolerance over mid-query faults. There is a wide range of analytical tools meant for processing the data. Most of them do not innately support distributed data. For instance open source based Pandas [12], library developed using Python offers high performance merging, joining and integrated indexing on the given dataset but not on distributed data.

RDBMS is good at handling structured data, but today's world is social media centric and technologies like web2.0 had enabled the Web to be read-write. Such advancements paved a way for more unstructured data. Hadoop and related ecosystem tools play a vigorous role in handling such situations. There are many graph database-related tools available such as Neo4j, Graphx, Pragel, Giraph, OrientDB, and GraphBuilder. All of these tools are open source, and each of them has its own capabilities. Following is an illustration of the proposed framework (Fig. 4).

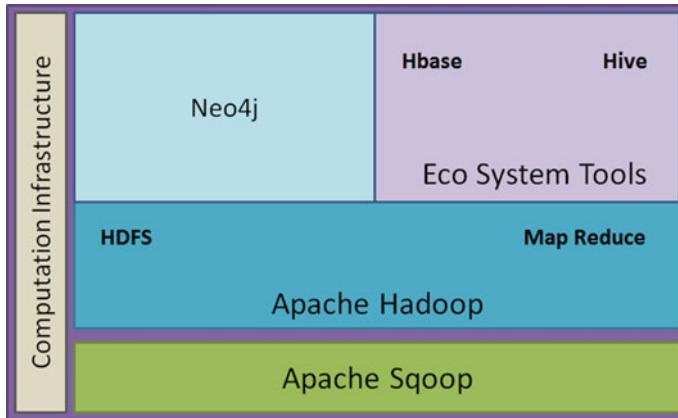


Fig. 4 Proposed framework

2.1 Apache Hadoop

Apache Hadoop is an open-source-based tool for managing the data across distributed nodes [13]. It contains two core components namely, HDFS and Map Reduce. HDFS—distributed file system, which coordinates the nodes for data storage, data consistency. Map Reduce is the programming extension available in Hadoop. Map Reduce offers parallel processing capabilities. Any task that is to be computed is executed as a Map Reduce job. An inherent feature of Hadoop is that it can run on commodity hardware.

2.2 Apache Sqoop, Hive, and Hbase

Apache Sqoop is a tool to efficiently move data between Hadoop and structured data stores like RDBMS. Sqoop supports most of the RDBMS such as MySQL and Postgresql. [14]. Hive is a data warehouse tool. It facilitates managing large datasets residing in distributed storage. Hive provides a mechanism to query the data using HiveQL. Also, it offers flexibility to map/reduce programmers to plug in their custom mappers and reducers instead of using HiveQL [15]. Apache Hbase is column-oriented database that runs on Hadoop. It organizes the large data on column-oriented storage. All the related information is managed through column families. Hbase supports millions of columns and billions of rows. Hbase leverages us to group all the information to be stored in a table, no matter how large is the table size [16].

2.3 Neo4j

Neo4j is an open-source graph database, developed using Java. Neo4j is an embedded, disk-based, fully transactional Java persistence engine. Neo4j uses graphs as primary data structure to store data. It stores everything in the form of either an edge, a node or an attribute. Every node and edge can have as many attributes as possible. Nodes and edges can be labelled [17, 18]. Neo4j supports bulk import of data from csv format; hence, the data from legacy systems are loaded into the Neo4j system for further analytics. Neo4j has Cypher, which is a declarative language similar to SQL for expressing patterns in graphs.

3 Related Work

3.1 Health Informatics System

Healthcare data from an existing application meant for managing health care are utilized to appraise the proposed framework. This application is built using J2EE technologies such as JSP, Servlets, and Hibernate. Postgresql, a popular open-source database, is used for storing the data. This application interfaces with the sensors which are meant for sensing various human body parameters such as heartbeat and blood pressure. Also, it has user interfaces meant for physicians, medication prescribed by them. All the data captured through the sensors are internal to this application; hence, Web services are provided to leverage the data exchange and interoperability. As the data stored in the database increase, performance of the system started decreasing. When the dataset is of large volume (in terms of TBs), analyzing the data stored in the RDBMS is tedious as it is computation intensive. It is attributed due to the increasing number of joins that is required among the tables.

3.2 Test Environment

Test environment is configured with six machines, and each of them is equipped with Intel i7 processor (4th generation) and 16 GB DDR3 RAM. Each machine has 1-TB secondary storage. All the machines are installed with 64 bit Linux (Ubuntu) operating system and Java 1.7.x 64 bit version. Out of these, one machine is configured as the Name Node and four machines are configured as the Data Nodes, and on another system, Neo4j is installed. In order to evaluate the proposed framework, synthetic data based on the existing application's data are generated by custom-made script. The test dataset size ranges from 100 GB to 2 TB.

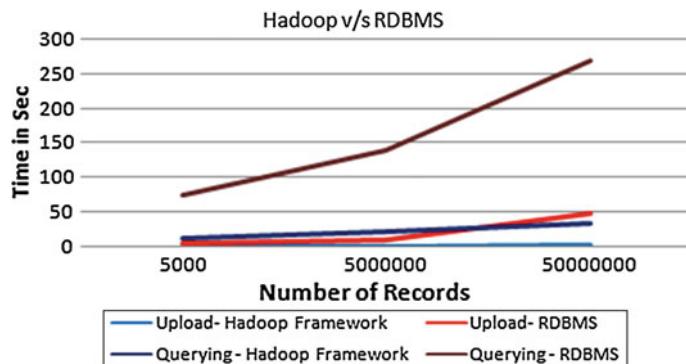


Fig. 5 Uploading and querying of records

The first analysis that is performed is to evaluate the performance of RDBMS with that of Hadoop tools. For this, basic operations such as upload data and query data are performed on records ranging from 5,000 to 50 million. Based on the empirical analysis conducted, the results were recorded and shown in Fig. 5.

As the dataset size increases, Hadoop tools perform much better than that of RDBMS. Also, Hadoop tools offers better fault tolerance measures over the RDBMS system by distributing the data and with the replication mechanism. Also, RDBMS-based system is prone to mid-query faults, whereas Hadoop-based framework does not have such limitation.

Subsequently, finding connected data among the large volume of data using RDBMS and Graph Database is conducted using the framework. Connected data in the healthcare data get a list of patients having same syndrome, along with the medication possibly with the expert physicians, who can diagnose that particular syndrome. RDBMS maintains links across the entities using primary key-foreign key link. An analysis of finding connected entities (nested linkage of entities) up to

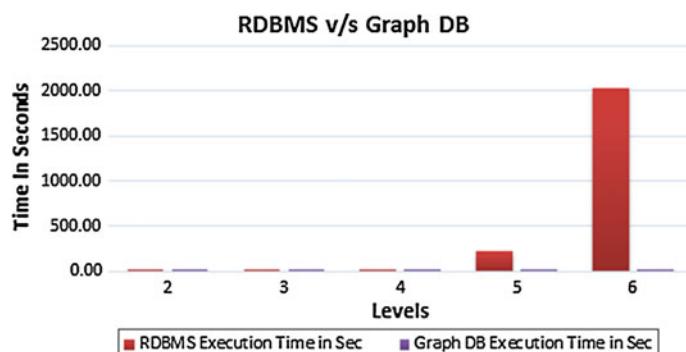


Fig. 6 RDBMS versus graph DB

six levels is conducted by considering 5,000 patients. Figure 6 is the outcome of the analysis performed.

From this, it is clear that Graph database performs better than the RDBMS. While the levels are up to four, there is no much difference in terms of performance is evidenced. Once the level goes beyond 5, RDBMS takes longer time. It is because of nested links, and as the levels increase, RDBMS needs more joins to compute the result, and it discards previous results when making more joins. Since each join results in Cartesian product, the complexity of such execution is in exponential order. Whereas, Neo4j maintains a list of previously visited nodes, even though the levels increase still it is able offer the same performance.

4 Conclusion

In this paper, we have elaborated the health informatics systems and their capabilities. We highlighted a data management framework for large volume of healthcare data and the capabilities of Graph Databases. Graph databases as explained by the proposed framework is suited well for the healthcare data to add smartness to the data itself by amalgamating connected data. Future enhancement to this framework is to add the machine learning mechanisms.

References

1. Koppel, R., Lehmann, C.U.: Implications of an emerging EHR monoculture for hospitals and healthcare systems. doi: [10.1136/amiajnl-2014-003023](https://doi.org/10.1136/amiajnl-2014-003023) pii: amiajnl-2014-003023 (2014)
2. <https://www.healthypeople.gov/2020/topics-objectives/topic/health-communication-and-health-information-technology>
3. Darwish, A., Hassanien, A.E.: Wearable and implantable wireless sensor network solutions for healthcare monitoring. Sensors **11**(6), 5561–5595 (2011)
4. <http://www.medscape.com/viewarticle/409940>
5. http://wiki.hl7.org/index.php?title=Main_Page
6. <http://www.hhs.gov/ocr/privacy/hipaa/understanding/>
7. <http://neo4j.com/developer/graph-db-vs-rdbms/>
8. Bao, Y., Ren, L., Zhang, L., Zhang, X., Luo, Y.: Massive sensor data management framework in Cloud manufacturing based on Hadoop. In: 10th IEEE International Conference on Industrial Informatics (INDIN), pp. 397–401 (2012)
9. Dean, J., Ghemawat, S.: MapReduce: simplified data processing on large clusters. In: Proceedings of the 6th Conference on Symposium on Operating Systems Design and Implementation, vol. 6, pp. 10 (2004)
10. Shvachko, K., Kuang, H., Radia, S., Chansler, R.: The Hadoop distributed file system. In: 2010 IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST), pp. 1–10 (2010)
11. <http://neo4j.com/developer/graph-db-vs-nosql/>
12. <http://pandas.pydata.org/>
13. <http://hadoop.apache.org/>

14. <http://sqoop.apache.org/>
15. <http://docs.treasuredata.com/articles/hive>
16. <http://hbase.apache.org/>
17. Dominguez-Sal, D., Urbón-Bayes, P., Giménez-Vañó, A., Gómez-Villamor, S., Martínez-Bazán, N., Larriba-Pey, J.L.: Survey of graph database performance on the HPC scalable graph analysis benchmark. Lecture Notes in Computer Science, vol. 6185, pp. 37–48 (2010)
18. <http://neo4j.com/developer/get-started/>

Capacity Analysis of ZF Receiver in Correlated MIMO Nakagami-m Channel

Samarendra Nath Sur, Soumyasree Bera, Rabindranath Bera and Bansibadan Maji

Abstract We investigate the ergodic capacity of multiple-input multiple-output (MIMO) system in Nakagami-m fading channels with linear receivers. Particularly in this paper, zero-forcing (ZF) receiver is considered. In this paper, we were able to demonstrate the impact of the channel impairment on the channel eigenvalues of the MIMO system and also represent the useful insights into the impact of the correlated Nakagami-m channel parameters (e.g., correlation coefficient and Nakagami fading parameter m).

Keywords MIMO · Nakagami-m · Correlated channel · Eigenvalue · Probability distribution function (PDF) · Cumulative distribution function (CDF)

1 Introduction

Multiple-input multiple-output (MIMO) antenna systems have gained substantial research interest as it provides high data throughput and significant enhancement in link reliability over single-antenna technology, without requiring extra power or bandwidth [1, 2]. From the point of view of practical implementation in multiple-input multiple-output spatial multiplexing (MIMO-SM) systems, linear zero forcing (ZF) receivers are important due to their minimal complexity requirements.

Lots of work has been done and also is going on to understand the capacity limits of MIMO systems in various practical propagation environments [3–5]. In order to understand the effects of random fading on the performance of a MIMO

S.N. Sur (✉) · S. Bera · R. Bera
Department of Electronics and Communication Engineering,
Sikkim Manipal Institute of Technology, Sikkim, India
e-mail: samar.sur@gmail.com

B. Maji
Department of Electronics and Communication Engineering,
National Institute of Technology, Durgapur, West Bengal, India

communication system, it is vital to understand the statistical behavior of the eigenvalues of $\mathbf{H}\mathbf{H}^*$ (\mathbf{H} represent the channel matrix). Among other things, the distribution of the channel eigenvalues determines the capacity of the system [6].

This paper deals with the Nakagami-m channel as it represents a more general fading model, which includes one-sided Gaussian fading, Rician fading, and Rayleigh fading model [7]. Therefore, analyzing the ergodic capacity of MIMO Nakagami-m fading channels is highly importance. In the recent past, some work has been done on this issue [3], [8–11]. We consider Vertical Bell Laboratories Layered Space–Time (V-BLAST) systems in correlated Nakagami-m fading channels with equal power allocated to each transmit antenna and also we consider that the channel state information (CSI) is available only at the receiver (CSIR).

This paper is arranged as follows: Sect. 2 presents the mathematical model for Nakagami-m channel, channel capacity calculation, and capacity difference for uncorrelated and correlated MIMO channel for ZF receiver. All the results are presented in Sect. 3.

2 Mathematical Model

2.1 Channel Model

We consider a MIMO system [11], with N_t transmitting antenna and N_r receiving antennas. The received signal y can be described by

$$y = \mathbf{H} x + n \quad (1)$$

where the transmit symbols vector x satisfies $E\left\{\|x\|^2\right\} \leq P$ (P is the total power), and n is the $N_r \times 1$ additive white Gaussian noise vector. The vector \mathbf{H} represents the correlated channels for the wireless transmission. The channel is assumed to follow a Nakagami-m fading probability distribution function (PDF). Let γ represents the instantaneous SNR, and it can be defined as follows:

$$\gamma = \beta^2 \frac{E_s}{N_0} \quad (2)$$

where β is the fading amplitude, E_s is the energy per symbol, and N_0 is the noise spectral density. The PDF of fading amplitude β for the Nakagami-m fading channel is given by

$$P_\beta(\beta) = \frac{2}{\Gamma(m)} \left(\frac{m}{\Omega}\right)^m \beta^{2m-1} \exp\left(-\frac{m\beta^2}{\Omega}\right), \quad \beta \geq 0, \quad m \geq \frac{1}{2} \quad (3)$$

where $\Gamma(\cdot)$ represents the gamma function, $\Omega = E[\beta^2]$ and m is the parameter of fading depth that ranges from 0.5 to infinity, and this parameter is responsible for the variation in fading condition. The fading parameter m is defined by the equation as given below

$$m = \frac{E^2[\beta^2]}{\text{var}[\beta^2]} \quad (4)$$

Then, the PDF of the instantaneous SNR γ is given by [12]

$$P_\gamma(\gamma) = \frac{1}{\Gamma(m)} \left(\frac{m}{\bar{\gamma}} \right)^m \gamma^{m-1} \exp\left(-\frac{m\gamma}{\bar{\gamma}}\right), \quad \gamma \geq 0 \quad (5)$$

2.2 Channel Capacity Calculation

In the paper we assume that CSI is known perfectly at the receiver and equal power P is distributed throughout the transmitter antennas.

The ergodic capacity of a MIMO system is given by,

$$C = E \left[\log_2 \det \left(I + \frac{P}{N_0 N_t} \mathbf{H} \mathbf{H}^\dagger \right) \right] \quad (6)$$

For convenience, we define $s = \min\{N_t, N_r\}$ and $k = \max\{N_t, N_r\}$ and

$$W = \begin{cases} \mathbf{H} \mathbf{H}^\dagger & N_r < N_t \\ \mathbf{H}^\dagger \mathbf{H} & N_r \geq N_t \end{cases} \quad (7)$$

The capacity can be written in terms of the eigenvalues $\lambda_1 \dots \lambda_n$ of W as follows:

$$C = E \left[\sum_{i=1}^n \log_2 \left(I + \frac{P}{N_0 N_t} \lambda_i \right) \right] \quad (8)$$

The channel is assumed to be independent and identically distributed (i.i.d) which follows a Nakagami-m fading PDF. As in [13], the ergodic capacity of a MIMO system in Nakagami-m channel can be represented in terms of Meijer G-function [14] as

$$C = \frac{s}{\Gamma(km) \ln 2} G_{3,2}^{1,3} \left(\frac{P}{N_0 N_t} \frac{\Omega}{m} \middle|^{1-km,1,1} \right) \quad (9)$$

And also as in [13], at high-SNR condition, we have

$$C_{\text{hsnr}} = s \log_2 \left(\frac{P}{N_t N_0} \right) + \frac{s}{\ln 2} \left[\psi(\text{km}) - \ln \left(\frac{m}{\Omega} \right) \right] \quad (10)$$

where $\psi(\cdot)$ is the digamma function.

From the above equation it is clear that the channel capacity increases linearly with the increase in minimum number on antennas. Now increase in transmitter antennas actually does not provide any capacity gain but increase in receiver antennas increases the ergodic capacity logarithmically as given below ($\Omega = 1$)

$$C_{\text{hsnr}}|_{N_r \rightarrow \infty} = N_t \log_2 \left(\frac{P}{N_t N_0} \right) + N_t \log_2 N_r \quad (11)$$

And also, if we increase both the transmitter and the receiver antennas, the channel ergodic capacity becomes a function of minimum number of antenna and $\log_2 \left(\frac{P}{N_0} \right)$.

From this, one important conclusion can be drawn that increase in number of antennas effectively reduces the effect of Nakagami fading parameter m. And also for Nakagami-m fading channels, in the low-SNR regime, the ergodic capacity can be expressed as follows:

$$C_{\text{lsnr}} = \frac{N_r * \text{SNR}}{\ln 2} \quad (12)$$

2.3 Capacity Difference for Uncorrelated and Correlated Channel

The channel capacities [5, 10] of a MIMO system with ZF receiver in uncorrelated and correlated channel condition are given below.

$$C_{\text{ZF}_{\text{UNCor}}} = \log_2 \left[1 + \frac{\text{SNR}}{(\mathbf{H}^H \mathbf{H})^{-1}} \right] = \log_2 [1 + \text{SNR} (\mathbf{H}^H \mathbf{H})] \quad (13)$$

$$C_{\text{ZF}_{\text{Cor}}} = \log_2 \left[1 + \frac{\text{SNR}}{(\mathbf{H}^H \mathbf{H} \mathbf{R}_{\text{RX}} \mathbf{R}_{\text{TX}})^{-1}} \right] = \log_2 [1 + \text{SNR} (\mathbf{H}^H \mathbf{H} \mathbf{R}_{\text{RX}} \mathbf{R}_{\text{TX}})] \quad (14)$$

$$\text{Now } \log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots \infty$$

$$C_{\text{ZF}_{\text{UNCor}}} = [\text{SNR}(H^H H)] \left[1 - \frac{\text{SNR}(H^H H)}{2} + O(\text{SNR}^2) \dots \infty \right] \quad (15)$$

$$C_{\text{ZF}_{\text{Cor}}} = [\text{SNR}(H^H H R_{\text{RX}} R_{\text{TX}})] * \left[1 - \frac{\text{SNR}(H^H H R_{\text{RX}} R_{\text{TX}})}{2} + O(\text{SNR}^2) \dots \infty \right] \quad (16)$$

Therefore,

$$\frac{C_{\text{ZF}_{\text{Cor}}}}{C_{\text{ZF}_{\text{UNCor}}}} = \frac{R_{\text{RX}} R_{\text{TX}} \left[1 - \frac{\text{SNR}(H^H H R_{\text{RX}} R_{\text{TX}})}{2} + O(\text{SNR}^2) \dots \infty \right]}{\left[1 - \frac{\text{SNR}(H^H H)}{2} + O(\text{SNR}^2) \dots \infty \right]}$$

= 1 for very low correlated channel

< 1 for fully correlated channel.

(As $\det(R_{\text{RX}} \text{ or } R_{\text{TX}})$ approaches to unity with the decrease in correlation coefficient). Above equation shows that with the increase in correlation coefficient, the capacity reduces.

3 Result

The eigenvalue analysis of a MIMO channel plays an important role to properly understand the throughput preformation of a MIMO system. Figures 1 and 2 represent the impact of the Nakagami fading parameter m and correlation coefficient over the distribution of the eigenvalues. With the increase in the correlation coefficient, independency in the paths is lost, and this in turn reduces the diversity order of the MIMO system. This reduction in the diversity order leads to lose in the throughput performance.

Figures 3, 4, and 5 represent the performance analysis of MIMO system in Nakagami- m channel with the variation in the correlation coefficient, with the variation in the number of antennas, and with the variation in the Nakagami fading parameter m , respectively. From Figs. 4 and 5, it is clear that increase in the correlation coefficient reduces the MIMO diversity order and thus reduces the throughput performance of the system. And also, with the variation in the m value, the channel statistic changes from one-sided Gaussian distribution to the Rician distribution. For $m = 0.5$, the PDF is one-sided Gaussian, whereas for $m = 1$, it is Rayleigh fading, and for $m = 2$, it is Rician fading. Figure 5 shows the variation in

Fig. 1 CDF of eigenvalues of 2×2 MIMO Nakagami-m channel (with the variation in m)

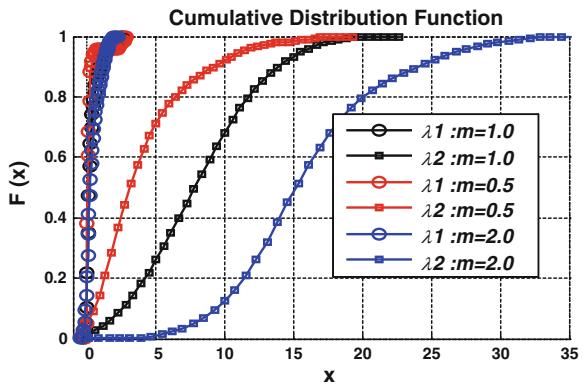


Fig. 2 CDF of eigenvalues of 2×2 MIMO Nakagami-m channel (with the variation in correlation coefficient)

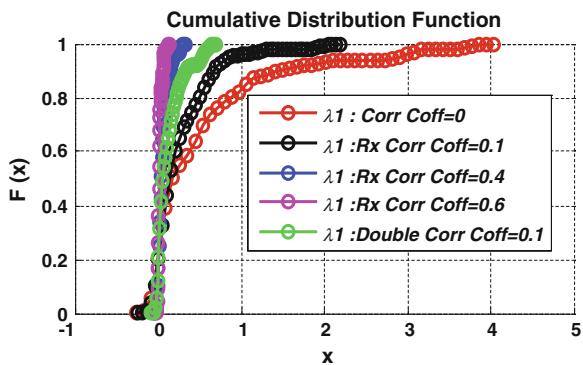


Fig. 3 Capacity of 4×4 MIMO system with Nakagami-m channel

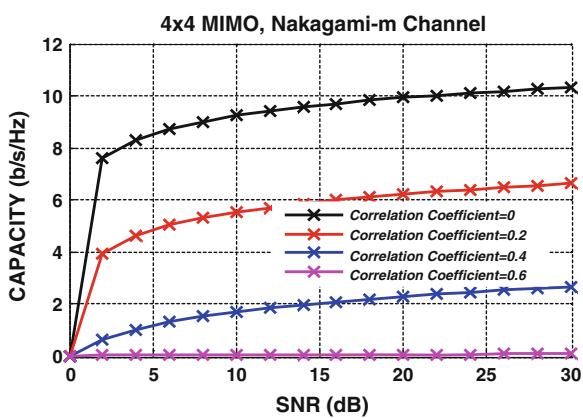


Fig. 4 Capacity of MIMO systems with Nakagami-m channel

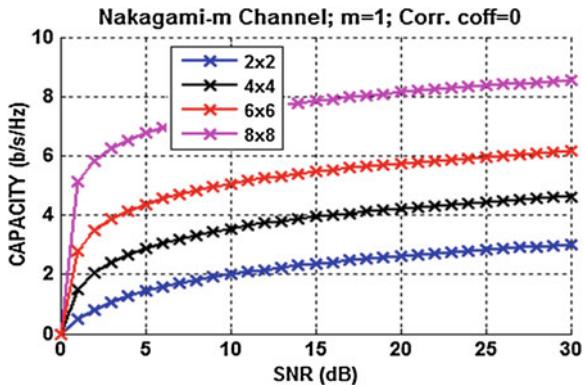


Fig. 5 Capacity of 8×8 MIMO system with Nakagami-m channel

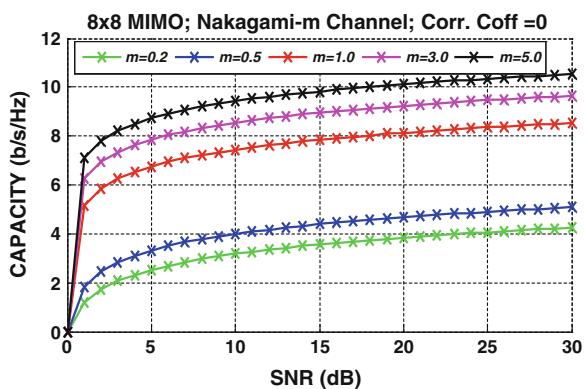
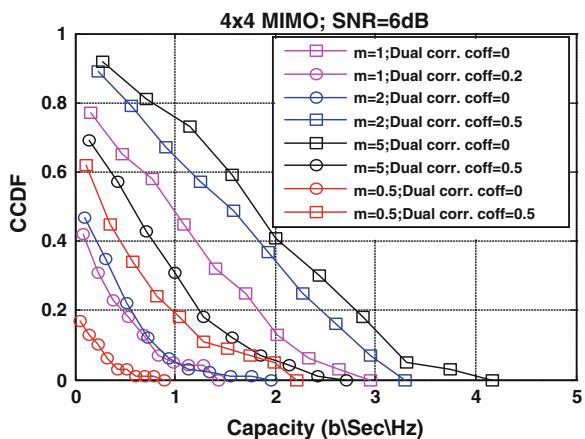


Fig. 6 CCDF of capacity of 4×4 MIMO system with Nakagami-m channel



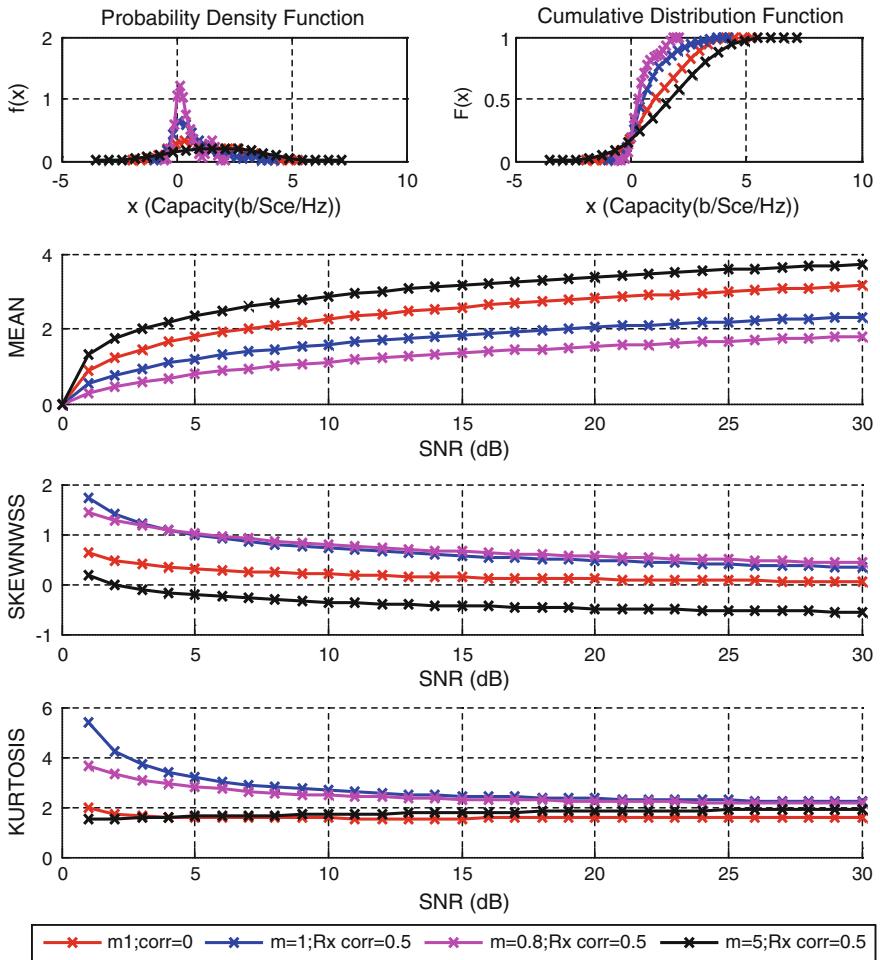


Fig. 7 Mean, variance, skewness, and kurtosis of the capacity (bits/s/Hz) versus SNR for Nakagami- m and doubly correlated (Rx corr = 0.5) 4×4 MIMO channels

channel capacity with the variation in m values. Figure 6 represents the CCDF variation in MIMO system with the variation in m values and correlation coefficient (Fig. 7).

Figure 7 shows the analytical and simulated mean, variance, skewness, and kurtosis of capacity for correlated Nakagami- m MIMO channels as a function of correlation coefficient and Nakagami fading parameter m . It can be seen that with the increase in SNR, the skewness and kurtosis start converging.

4 Conclusion

In this paper, we have derived an expression for capacity in correlated ZF receiver. This paper deals with the capacity analysis of MIMO system in Nakagami-m channel with the variation in correlation coefficient and Nakagami-m fading parameter. As in the simulated results, the correlation coefficient and Nakagami shape parameter m have significant influence over the throughput of the MIMO system.

References

1. Telatar, I.E.: Capacity of multi-antenna Gaussian channels. *Europ. Trans. Telecommun.* **10**(6), 585–595 (1999)
2. Foschini, G.J., Gans, M.J.: On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Pers. Commun.* **6**, 311–335 (1998)
3. Zhong, C., Matthaiou, M., Huang, A., Zhang, Z.: On the sum rate of MIMO Nakagami-m fading channels with linear receivers. *IEEE Trans. Wireless Commun.* **11**(10), 3651–3659 (2012)
4. Shin, H., Lee, J.H.: Capacity of multi-antenna fading channels: spatial fading correlation, double scattering, and keyhole. *IEEE Trans. Inf. Theory* **49**(10), 2636–2647 (2003)
5. Sur, S.N., Ghosh, D., Bhaskar, D., Bera, R.: Contemporary MMSE and ZF receiver for V BLAST MIMO system in Nakagami-m fading channel. In: IEEE INDICON (2011), pp. 1–5, 16–18 Dec 2011
6. Martin, C., Ottersten, B.: Asymptotic eigenvalue distributions and capacity for MIMO channels under correlated fading. *IEEE Trans. Wireless Commun.* **3**(4), 1350–1359 (2004)
7. Nakagami, M.: The m-distribution—a general formula of intensity distribution of rapid fading. In: Hoffman, W.C. (ed.), *Statistical methods in radio wave propagation*, pp. 3–36, Pergamon (1960)
8. Musavian, L., Aïssa, S.: On the achievable sum-rate of correlated MIMO multiple access channel with imperfect channel estimation. *IEEE Trans. Wireless Commun.* **7**(7), 2549–2559 (2008)
9. Hanlen, L., Grant, A.: Capacity analysis of correlated MIMO channels. *IEEE Trans. Inf. Theory* **58**(11), 6773–6787 (2012)
10. Sur, S.N., Bera, R., Maji, B.: Channel capacity analysis of MIMO system in correlated Nakagami-m fading environment. *Int. J. Eng. Trends Technol. (IJETT)* **9**(3), 101–105 (2014)
11. Sur, S.N., Ghosh, D.: Channel capacity and BER performance analysis of MIMO system with linear receiver in Nakagami channel. *Int. J. Wireless Microwave Technol* **3**(1), 26–36 (2013)
12. Shin, Hyundong, Lee, JaeHong: On the error probability of binary and M-ary signals in Nakagami-m fading channels. *IEEE Trans. Commun.* **52**(4), 536–539 (2004)
13. Zhong, C., Wong, K.-K., Jin, S.: Capacity bounds for MIMO Nakagami-m fading channels. *IEEE Trans. Signal Process.* **57**(9), 3613–3623 (2009)
14. Prudnikov, A.P., Brychkov, Y.A., Marichev, O.I.: *Integrals and Series: More Special Functions*, vol. 3. Gordon and Breach, New York (1990)

Neural Estimations of First Null Beamwidth for Broadside and End-Fire Uniform Linear Antenna Arrays

Subhash Mishra, Ram Narayan Yadav and Rajendra Prasad Singh

Abstract Antenna arrays are capable of reducing co-channel interference and multipath fading. Beamwidth parameters have key role in the performance of an antenna array. Estimation of the first null beamwidth (FNBW) is important for the design of the array. Artificial neural networks (ANNs), also known as neural networks (NNs), use simple mathematical tools. The ability of trained ANNs to predict results for the unseen inputs makes them suitable for real-time applications. They can map the nonlinear behaviour of antenna arrays easily. This paper presents the neural estimations of the FNBW parameter for the broadside and end-fire uniform linear antenna arrays (ULAs), using radial basis function neural networks (RBF-NNs). Precise estimation of FNBW helps in achieving desired accuracy in array design and operation.

Keywords Antenna arrays · Uniform linear antenna arrays (ULAs) · First null beamwidth (FNBW) · Artificial neural networks (ANNs) · Radial basis function neural networks (RBF-NNs)

1 Introduction

An antenna array is the assembly of multiple antenna elements arranged in a suitable configuration to produce the desired radiation characteristics. The multiplicity of elements helps in achieving desired patterns and shapes [1, 2]. Traditional

S. Mishra (✉) · R.N. Yadav · R.P. Singh

Department of Electronics and Communication Engineering, Maulana Azad National Institute of Technology (MANIT), Bhopal, India
e-mail: subhashmishra67@gmail.com

R.N. Yadav
e-mail: rnyadav@gmail.com

R.P. Singh
e-mail: prof.rpsingh@gmail.com

means of dealing with the factors of impairments which affect the coverage and capacity of wireless networks are not sufficient. In such a scenario, antenna arrays have emerged as highly effective tools to address these problems related to the performance of wireless networks [3, 4]. Using powerful signal processing and space-division multiple access (SDMA), antenna arrays provide ‘smart antennas’ which can solve various problems of wireless networks. Smart antenna arrays help in reducing the signal impairments such as co-channel interference and multipath fading. They provide improvement in the spectrum utilization and network capacity [4, 5].

A broadside linear array of isotropic elements has a wide radiation pattern in the plane towards the normal of the array axis and has a narrow pattern in the plane which includes the array axis. On the other hand, the radiation pattern of an end-fire array is along the axis. For linear arrays, the shape of the beam changes from ‘disc’ beam (at the broadside) to ‘pencil’ beam (at the end-fire) [1, 2, 6].

Beamwidth is a figure of merit in the performance of an antenna array and is defined as the angular separation or the width between two identical points on the opposite sides of the pattern maximum [1]. A small beamwidth would mean a large aperture and vice versa. The first null beamwidth (FNBW) and the half-power beamwidth (HPBW) both are significant in the working of antenna arrays. The FNBW has important role in the operation of an antenna array and needs to be considered in the design of the array [1–3, 6].

Artificial neural networks (ANNs) use basic mathematical tools and can provide results faster than many conventional methods. They can deal with the nonlinear properties of the arrays easily as they are also nonlinear in nature [7, 8]. In this study, we have estimated the FNBW for 8-element and 16-element broadside and end-fire uniform linear arrays (ULAs) using radial basis function neural networks (RBF-NNs) which are well known for their faster training and remarkable interpolation abilities [9–12]. The organization of the remaining part of the paper is as follows. Section 2 covers the FNBW expressions for the broadside and the end-fire arrays. Section 3 covers a brief overview on the neural networks (NNs) and the NN model used in this paper. Simulation results are presented in Sect. 4. Section 5 concludes the paper.

2 First Null Beamwidth (FNBW)

The null is a point in the radiation pattern of an antenna array, where the response of the array is zero. The angular separation between the first nulls on either side of the pattern is referred to as the FNBW θ_{FN} . This is also known as the null-to-null beamwidth (NNBW). It has an important role in array design [1, 3].

2.1 FNBW for Broadside Arrays

For a broadside uniform linear array with M antenna elements on z -axis, and inter element spacing d , the FNBW ($\theta_{\text{FN-BS}}$) is expressed as follows [1, 3]:

$$\theta_{\text{FN-BS}} = 2 \left[\frac{\pi}{2} - \cos^{-1} \left(\frac{\lambda}{Md} \right) \right] \quad (1)$$

2.2 FNBW for End-Fire Arrays

The FNBW ($\theta_{\text{FN-EF}}$) for an end-fire uniform linear array is expressed as follows [1, 3]:

$$\theta_{\text{FN-EF}} = 2 \cos^{-1} \left[1 - \left(\frac{\lambda}{Md} \right) \right] \quad (2)$$

3 Artificial Neural Networks

ANNs are nonlinear in nature, and therefore, they can handle the nonlinear behaviour of the antenna array elements properly. They have massive parallel processing potential and the ability to provide fast solutions. The concept of ANNs (or simply NNs) is developed from neurons of the human brain [7, 8]. ANNs also learn from examples, and after learning (also known as training), they can interpolate results accurately even for the unseen inputs [9, 10]. Because of these reasons, NNs are very popular for solving the problems related to antenna arrays [11–13]. Some reviews on the use of NNs in antenna array design, synthesis, analysis, estimation of direction of arrival (DOA) and beam formation applications can be found in [14–16].

It can be seen from Eqs. (1) and (2) that the FNBW θ_{FN} depends on the number of elements M and the inter element spacing d . The NN model used for predicting the FNBW for broadside and end-fire ULAs is shown in Fig. 1. The NN model is able to predict the output after obtaining proper training. The training data for the NN is chosen from the theoretical values of input and output using Eqs. (1) and (2). After completion of training, the testing is performed on the data which has not been used in training.

RBF-NNs have been used in the study for the prediction of FNBW. The RBF-NN consists of three layers: an input layer, an output layer and a hidden layer. A neuron in a layer is connected to each neuron of the next layer through synaptic weights. Neurons belonging to the same layer are not connected with each other [7].

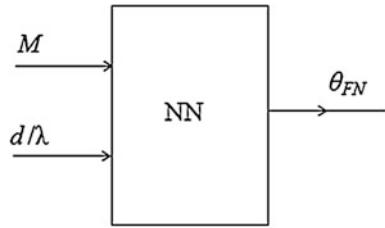


Fig. 1 Neural network model for FNBW estimation

The activation function for the neurons of the input and output layers is usually linear, whereas for the hidden layer, the activation function is radial basis transfer function [7–11] which follows the characteristics of a Gaussian function as follows:

$$\varphi(\|x - x_i\|) = \exp\left(\frac{-\|x - x_i\|^2}{2\sigma^2}\right) \quad (3)$$

where, σ is the spread parameter and x_i is the i th centre vector of radial basis function.

The output for the network can be expressed as follows [7–11]:

$$y = F(x) = \sum_{i=1}^P w_i \varphi(\|x - x_i\|) \quad (4)$$

where, x and y are the input and output of the network and w_i is the i th weight vector.

4 Simulation Results

Simulations are performed on 8-element and 16-element (isotropic) broadside and end-fire uniform linear antenna arrays (ULAs). The wavelength λ considered in the study is 12.5 cm. RBF-NNs using Levenberg–Marquardt (LM) algorithm have been used for estimating the FNBW parameter. A mean square error (MSE) of 10^{-6} is set as the performance goal. The spread constant has been considered as 0.5.

The RBF-NN for the estimation of FNBW for 8-element broadside array requires 35 neurons in the hidden layer, whereas for the 8-element end-fire array, an RBF-NN with 23 neurons in the hidden layer is found suitable.

The training data (input–output pairs) are obtained using Eqs. (1) and (2) as mentioned earlier. The training of each RBF-NN is performed with 80 % of data, and the remaining 20 % data, which is not included in training, is used in testing. The performance plot for the RBF-NN used for the FNBW estimation of the 8-element broadside array is shown in Fig. 2.

Fig. 2 RBF-NN performance
(8-element broadside ULA)

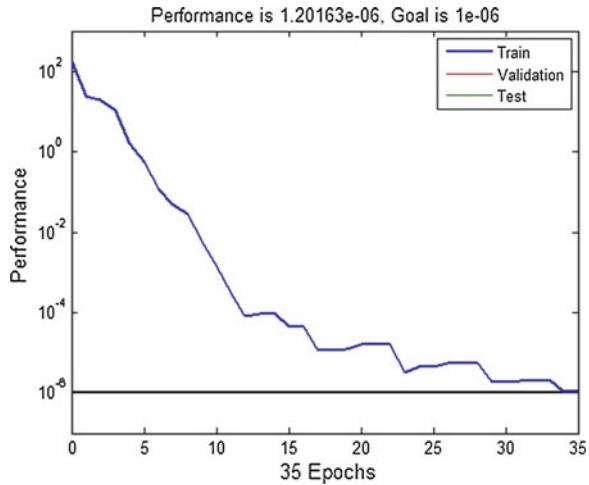
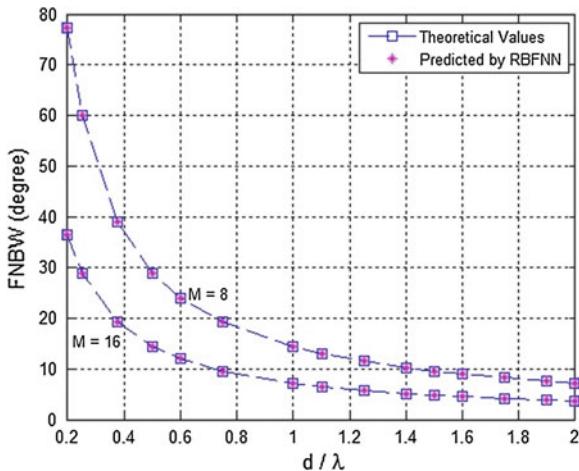


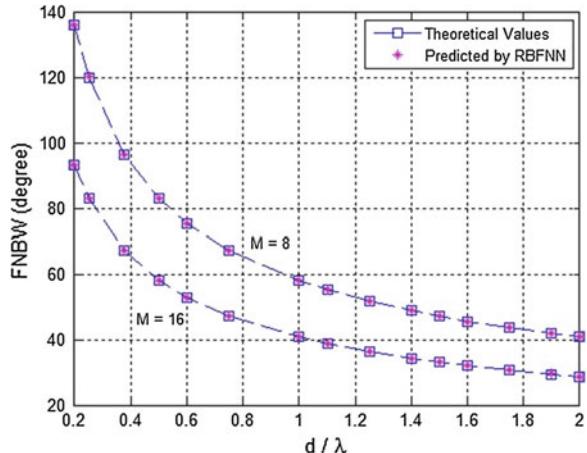
Fig. 3 FNBW plots for 8-element ($M = 8$) and 16-element ($M = 16$) broadside ULAs



The plots for the theoretical (desired) values and RBF-NN estimated values of FNBW as function of d/λ , for the 8-element ($M = 8$) and 16-element ($M = 16$) broadside ULAs are presented in Fig. 3. Similarly, the theoretical (desired) and the RBF-NN estimated values of FNBW as function of d/λ , for the 8-element and 16-element end-fire ULAs are shown in Fig. 4.

As evident from these plots, the estimated FNBW values are in agreement with the desired values. The average absolute errors (on the basis of 10 simulation runs for each case) in the estimations of FNBW for the 8-element and 16-element broadside arrays are 0.0210 and 0.0112 degrees, respectively. These errors for the 8-element and 16-element end-fire arrays are 0.0102 and 0.0079 degrees, respectively.

Fig. 4 FNBW plots for 8-element ($M = 8$) and 16-element ($M = 16$) end-fire ULAs



The RBF-NNs for the 8-element and 16-element broadside arrays require average run-times of 8.63 and 5.08 s, respectively, whereas the average run-times required by the RBF-NNs for the 8-element and 16-element end-fire arrays are 6.38 and 4.72 s, respectively. The neural network toolbox in MATLAB 2012b [17] environment has been used in simulations, on a computer having Intel Core i3-3110 M CPU at 2.4 GHz and 4 GB RAM.

5 Conclusion

Neural estimates of FNBW for the 8-element and 16-element (isotropic) broadside and end-fire ULAs have been obtained using RBF-NNs trained with LM algorithm. The results provided by this neural method are very close to the desired values even for the unseen inputs in the data set. The RBF-NNs have shown remarkable accuracy and speed. Accurately estimated beamwidth values would ensure precision in the design and operation of antenna arrays.

Acknowledgements We express our sincere thanks to the authors and the publishers of all the books, research papers/articles/reviews referred in this paper.

References

1. Balanis, C.A.: Antenna Theory, Analysis and Design, 3rd edn. Wiley, New Jersey (2005)
2. Mailloux, R.J.: Phased Array Antenna Handbook, 2nd edn. Artech House, Norwood (2005)
3. El Zooghby, A.H.: Smart Antenna Engineering. Artech House, Norwood (2005)
4. Godara, L.C.: Applications of antenna arrays to mobile communications part 1: performance improvement, feasibility, and system considerations. Proc. IEEE **85**(7), 1031–1060 (1997)

5. Winters, J.H.: Smart antennas for wireless systems. *IEEE Pers. Commun.* **5**(1), 23–27 (1998)
6. Hansen, R.C.: *Phased Array Antenna*, 2nd edn. Wiley, New Jersey (2009)
7. Haykins, S.: *Neural Networks: A Comprehensive Foundation*, 2nd edn. Prentice Hall, New Jersey (1998)
8. Zhang, Q.J., Gupta, K.C.: *Neural Networks for RF and Microwave Design*. Artech House, Norwood (2000)
9. Lo, T., Leung, H., Litwa, J.: Radial basis function neural network for direction-of-arrivals estimation. *IEEE Signal Process. Lett.* **1**(2), 45–47 (1994)
10. Southall, H.L., Simmers, J.A., O'Donnell, T.H.: Direction finding in phased arrays with a neural network beamformer. *IEEE Trans. Antennas Propag.* **43**(12), 1369–1374 (1995)
11. El Zooghby, A.H., Christodoulou, C.G., Georgopoulos, M.: Performance of radial basis function networks for direction of arrival estimation with antenna arrays. *IEEE Trans. Antennas Propag.* **45**(11), 1611–1617 (1997)
12. Rama Sanjeev Reddy, B., Vakula, D., Sarma, N.V.S.N.: Design of multiple function antenna array using radial basis function neural network. *J. Microwaves, Optoelectron Electromagn. Appl.* **12**(1), 210–216 (2013)
13. Merad, L., Bendimerad, F.T., Meriah, S.M., Djennas, S.A.: Neural networks for synthesis and optimization of antenna arrays. *Radio Eng.* **16**(1), 23–30 (2007)
14. Patnaik, A., Anagnostou, D.E., Mishra, R.K., Cristodoulou, C.G., Lyke, J.C.: Applications of neural networks in wireless communications. *IEEE Antennas Propag. Mag.* **46**(3), 130–137 (2004)
15. Du, K.L., Lai, A.K.Y., Cheng, K.K.M., Swamy, M.N.S.: Neural methods for antenna array signal processing: a review. *Sig. Process.* **82**(4), 547–561 (2002)
16. Mishra, S., Yadav, R.N., Singh, R.P.: Neural networks in smart antenna design for co-channel interference (CCI) reduction: a review. In: Proceedings of 5th International Conference on Wireless and Mobile Communications (ICWMC), Cannes La Bocca, France, pp. 228–233, (2009)
17. Beale, M.H., Hagon, M.T., Demuth, B.: *Neural Network ToolboxTM*, User's Guide, R2012b. MATLAB, MathWorks, USA

Representing the Extended H_∞ Filter (EHF) as a Modified Extended Kalman Filter (EKF) for Predicting the Range of the Robustness Bound

Manika Saha, Ratna Ghosh and Bhawati Goswami

Abstract In this paper, the extended H_∞ filter (EHF) has been logically reformulated as a modification of the extended Kalman filter (EKF). Thereafter, a robustness metric and a sensitivity metric have been defined, analogous to that in Saha et al. (IEEE Trans Instrum Meas 63(4):964–971, 2014 [1]), for the EHF which is used to determine a suitable combination of the filter tuning parameters. The major objective of this paper is to predict the usable range of the robustness bound γ for obtaining reliable filter performances. For this purpose, the same problem as in Saha et al. (IEEE Trans Instrum Meas 63(4):964–971, 2014 [1]) of tracking a 2D ballistic target has been studied and the desired states of the EHF have been chosen to be the same as the measurements. The plots of the metrics have been obtained for various choices of γ . The corresponding RMSE performances have then been plotted for each of these choices of γ , while the filter process noise covariance has been varied over the range obtained from the metric plot. It is observed that the metrics are useful in determining the usable range of γ for which the designer is able to obtain the desired trade-off between robustness and sensitivity.

Keywords EHF · Bound · Robustness · Sensitivity · RMSE · Metrics · Filter tuning

M. Saha (✉)

Department of Electrical Engineering, Future Institute of Engineering and Management,
Sonarpur, Kolkata 700150, India
e-mail: manikasaha@gmail.com

R. Ghosh · B. Goswami

Department of Instrumentation and Electronics Engineering, Jadavpur University,
2nd Campus, Kolkata 700098, India
e-mail: rg@iee.jusl.ac.in

B. Goswami

e-mail: bg@iee.jusl.ac.in

1 Introduction

In certain applications, several researchers claim the robustness of the extended H_∞ filter (EHF) over the extended Kalman filter (EKF) [2]. These claims are based on the perceived advantages of the EHF over the EKF that it is not required to know the process or measurement noise statistics accurately; instead, the designer can specify certain desired outputs and corresponding robustness bounds for the filter based on apriori knowledge of the magnitudes of the noises and initial estimation error. An investigation into the interrelations of the two optimization approaches has established that the Kalman filter may be considered to be the H_∞ filter for the particular case when the desired outputs are all the states, and there is no finite bound on the robustness requirement of the filter. On the other hand, the H_∞ filter may be considered to be a robust version of the Kalman filter in which the filter process noise covariance Q is replaced by an effective covariance, say Q_{eff} , which is dependent on the choice of the desired outputs and robustness bound γ [3, 2]. However, a seamless derivation of the EHF as an extension of the EKF is still lacking for the general case.

On the other hand, robustness and sensitivity metrics have been derived for the EKF in [1] and these have been used to predict the robustness or sensitivity of the RMSE performances of the filter in estimating the states of a 2D falling body. The metrics have accordingly been used to obtain a suitable choice of the filter tuning parameters, specifically the filter process noise covariance. In view of the established correlations of the two approaches of the EKF and the EHF, it is thus valid to explore whether it is possible to provide a guideline similar to that for the EKF for choosing the filter tuning parameters of the EHF. In particular, since the EHF is designed to provide desired robustness, is it possible to identify a suitable range of the robustness bound and evaluate its effect on the robustness and sensitivity of the filter RMSE performances? This paper is organized as follows: The EHF algorithm is stated in Sect. 2 as a modification of the standard EKF algorithm. The definitions of the robustness and sensitivity metrics for the general EHF as well as the derivation of suitable metrics for the specific case, when the desired states are identically the measurements, are also stated in this Section. Thereafter, the problem statement is stated in Sect. 3. Section 4 contains the results and discussion, while the conclusions are stated in Sect. 5.

2 EHF Algorithm: Modified EKF Approach

Any practical discrete-time system may be described at a particular time instant k using general linearized state and observation equations as in the EKF. An additional equation for the desired output z_k is added for the EHF. The total system description for the EHF is thus as follows:

$$\begin{aligned} x_{k+1} &= F_k x_k + G_k u_k + w_k \\ y_k &= H_k x_k + v_k \\ z_k &= L_k x_k. \end{aligned} \quad (1)$$

where F_k and H_k are the Jacobians of the functions $f(x_k)$ and $h(x_k)$ as in (28) and (29), respectively. As in the standard EKF, the state vector x_k , the observation vector y_k , and the known input u_k have dimensions n , m , and p , respectively, while F_k , H_k , and G_k are the respective state transition, measurement, and input matrices. Both H_k and G_k are not constrained to be time invariant.

The desired output vector z_k of dimension r is specifically defined in the EHF by the user, and L_k is the corresponding output matrix of dimension $r \times n$. The r desired outputs, which may be any linear combination of the states x_k , are chosen by the designer to reflect the specific requirement of a particular application. Additionally, however w_k and v_k are the usual n -dimensional process disturbance and m -dimensional measurement disturbance, respectively, yet Q_k and R_k are not necessarily the noise covariances as in the EKF; instead, these are weighting matrices whose elements represent the desired weights/penalties determined by the user for the individual and/or interaction components of the corresponding disturbances w_k and v_k . However, for this application, these have been nominally chosen as the respective covariance matrices as considered in [1].

For the system description in (1), the standard energy cost function J for the H_∞ filter optimization [2] can be stated as

$$J = \frac{\sum_{k=0}^N \|z_k - L_k \hat{x}_k^+\|_{R_k^{-1}}^2}{\|x_0 - \hat{x}_0^+\|_{P_0^{-1}}^2 + \sum_{k=0}^N (\|w_k\|_{Q_k^{-1}}^2 + \|v_k\|_{R_k^{-1}}^2)} < \gamma \quad (2)$$

where γ and R_k represent the user-defined values of the robustness bound and a weighting matrix corresponding to the desired outputs, respectively, while all other terms are as defined earlier.

The overall system can further be represented in terms of an augmented output y_{ak} as

$$\begin{aligned} x_{k+1} &= F_k x_k + G_k u_k + w_k \\ y_{ak} &= [y_k^T \ z_k^T]^T = [H_k^T \ L_k^T]^T x_k + [I_m \ 0_r]^T v_k \\ &= H_{ak} x_k + \Gamma_{ak} v_k. \end{aligned} \quad (3)$$

Here, the augmented output y_{ak} includes the measured outputs y_k and the desired outputs z_k and H_{ak} denote the corresponding augmented output matrix, while the $(m+r) \times m$ matrix Γ_{ak} is defined to incorporate the measurement noise in the system. In the reformulation of the suboptimal H_∞ problem, Forssell [3] uses an augmented measurement weighting matrix $R_{e,i}$ which makes the difference between the EKF and EHF formulation. This diagonal matrix contains R_k as well as a new

term $(-\gamma^2 I)$ pertaining to the robustness cost as the diagonal elements. Accordingly, in the present paper also, the cost for robustness in terms of the user-defined values of γ and R_{k1} is incorporated into an augmented measurement weighting matrix R_{ak} which is defined as

$$R_{ak} = \begin{bmatrix} R_k & 0 \\ 0 & -\gamma R_{k1} \end{bmatrix}. \quad (4)$$

In the EHF algorithm stated hereafter, it is logical to utilize the innovation q_k , R_{ak} as in (4), and an equivalent filter gain, henceforth, defined in terms of another matrix M_{ak} and denoted as K_k^H . The notations \hat{x}_k^- and P_k^- denote the apriori state estimate and estimation error covariance, respectively, while q_k , S_k , K_k , and P_k^+ denote the innovation, innovation covariance, filter gain, and aposteriori state error covariance, respectively. Using these, the corresponding sequential EHF formulation for this system can be stated in a form analogous to that stated in [2] as

$$\hat{x}_k^- = F_{k-1}\hat{x}_{k-1}^+ + G_{k-1}\hat{u}_{k-1} \quad (5)$$

$$P_k^- = F_{k-1}P_{k-1}^{H+}F_{k-1}^T + Q_{k-1} \quad (6)$$

$$q_k = y_k - H_k\hat{x}_k^- \quad (7)$$

$$M_{ak} = (I_n + H_{ak}^T R_{ak}^{-1} H_{ak} P_k^-)^{-1} \quad (8)$$

$$K_k^H = P_k^- M_{ak} H_k^T R_k^{-1} \quad (9)$$

$$\hat{x}_k^+ = \hat{x}_k^- + K_k^H q_k \quad (10)$$

$$P_k^{H+} = (I_n - K_k^H H_k) P_k^- = P_k^- M_{ak}. \quad (11)$$

2.1 Robustness and Sensitivity Metrics for EHF

Using this formulation of the EHF along with the condition to be satisfied, the instantaneous sensitivity and robustness metrics, J_{1ak} and J_{2ak} , respectively, for the general EHF may be stated analogous to those for the EKF stated in [1] as

$$J_{2ak} = \text{tr} \{(A_{ak} + B_{ak})^{-1} B_{ak}\} \quad (12)$$

and

$$J_{1ak} = \text{tr} \{(A_{ak} + B_{ak} + R_{ak})^{-1} R_{ak}\} \quad (13)$$

where

$$A_{ak} = H_{ak} F_{k-1} P_{k-1}^{H+} F_{k-1}^T H_{ak}^T, \quad (14)$$

and

$$B_{ak} = H_{ak} Q_{k-1} H_{ak}^T, \quad (15)$$

and the overall *augmented* sensitivity and robustness metrics, denoted as J_{1a} and J_{2a} , respectively, are defined in case of the EHF for the total time horizon N as

$$J_{1a} = \frac{1}{N} \sum_{k=1}^N J_{1ak} = \frac{1}{N} \sum_{k=1}^N \text{tr} \{ D_{ak} R_{ak} \} \quad (16)$$

$$J_{2a} = \frac{1}{N} \sum_{k=1}^N J_{2ak} = \frac{1}{N} \sum_{k=1}^N \text{tr} \{ (C_{ak})^{-1} B_{ak} \} \quad (17)$$

where

$$C_{ak} = (A_{ak} + B_{ak}) \quad (18)$$

and

$$D_{ak} = (A_{ak} + B_{ak} + R_{ak})^{-1}. \quad (19)$$

In order to provide a common basis for the comparison of the EKF and the EHF, the controlling parameter for the EHF metrics is considered to be the same as that for the EKF, which is $n_q = \log_{10} \left\{ \frac{1}{N} \sum_{k=1}^N \text{tr}(B_k) \right\}$ where $B_k = H_k Q_{k-1} H_k^T$.

2.1.1 A Special Case When $L_k = H_k$

As stated in Sect. 3, the desired states are assumed to be the same as the measurements, that is, $z_k = y_k$ and so $L_k = H_k$ in this paper. For this special case, the augmented robustness metric J_{2a} for EHF can be rewritten using block-partitioned matrices as

$$\begin{aligned} J_{2a} &= \frac{1}{N} \sum_{k=1}^N \text{tr} \{ (C_{ak})^{-1} B_{ak} \} \\ &= \frac{1}{N} \sum_{k=1}^N \text{tr} \begin{bmatrix} C_{ak}^{11} & C_{ak}^{12} \\ C_{ak}^{21} & C_{ak}^{22} \end{bmatrix}^{-1} \begin{bmatrix} B_{ak}^{11} & B_{ak}^{12} \\ B_{ak}^{21} & B_{ak}^{22} \end{bmatrix}. \end{aligned} \quad (20)$$

Since $L_k = H_k$, the matrix $C_{ak} = (A_{ak} + B_{ak})$ becomes singular and noninvertible, and all four block partitions of A_{ak} , B_{ak} , and C_{ak} are identical, with $A_{ak}^{ij} = H_k F_{k-1} P_{k-1}^{H+} F_{k-1}^T H_k^T$, $B_{ak}^{ij} = H_k Q_{k-1} H_k^T$ and $C_{ak}^{ij} = H_k P_k^- H_k^T$ for all $i, j = [1, 2]$. So, the nonsingular m -dimensional $(1, 1)$ block partitions of C_{ak} and B_{ak} are used to define another robustness metric $J_2^H = \frac{1}{N} \sum_{k=1}^N \text{tr} \{(C_{ak}^{11})^{-1} B_{ak}^{11}\} = \frac{1}{N} \sum_{k=1}^N \text{tr} \{(A_{ak}^{11} + B_{ak}^{11})^{-1} B_{ak}^{11}\}$.

On the other hand, an analysis of the augmented sensitivity metric J_{1a} for this particular case $L_k = H_k$ shows that there is no such singularity as in the case of J_{2a} . This can be ascribed to the presence of nonidentical block diagonal elements R_k and $-\gamma R_{k1}$ in the matrix R_{ak} . In the expression for D_{ak} as stated in (19), it is observed that this nonsingular matrix R_{ak} adds to the singular matrix C_{ak} . Thus, both R_{ak} and D_{ak} are nonsingular in this case.

However, since the robustness metric J_2^H to be considered in this case is m -dimensional, hence, a compatible m -dimensional sensitivity metric is derived in this case as $J_1^H = \frac{1}{N} \sum_{k=1}^N \text{tr} \{(D_{ak}^{11})R_k\}$. This is obtained as the $(1, 1)$ block-partitioned matrix of the augmented sensitivity metric J_{1a} which is

$$\begin{aligned} J_{1a} &= \frac{1}{N} \sum_{k=1}^N J_{1ak} = \frac{1}{N} \sum_{k=1}^N \text{tr} \{(D_{ak}^{11})R_{ak}\} \\ &= \frac{1}{N} \sum_{k=1}^N \text{tr} \begin{bmatrix} D_{ak}^{11} & D_{ak}^{12} \\ D_{ak}^{21} & D_{ak}^{22} \end{bmatrix} \begin{bmatrix} R_k & 0 \\ 0 & -\gamma R_{k1} \end{bmatrix}. \end{aligned} \quad (21)$$

Using the formula for inverses of block-partitioned matrices [4], an effective Q_k denoted as Q_{k-1}^{eff} , and the term D_{ak}^{11} can be evaluated as

$$\begin{aligned} Q_{k-1}^{\text{eff}} &= Q_{k-1} - (P_k^- H_k^T \{H_k P_k^- H_k^T \\ &\quad - (\gamma R_{k1})\}^{-1} H_k P_k^-), \end{aligned} \quad (22)$$

$$\begin{aligned} D_{ak}^{11} &= [\{C_{ak}^{11} + R_k\} - C_{ak}^{12} \{C_{ak}^{22} - (\gamma R_{k1})\}^{-1} C_{ak}^{21}]^{-1} \\ &= [A_{ak}^{11} + H_k Q_{k-1}^{\text{eff}} H_k^T + R_k]^{-1}. \end{aligned} \quad (23)$$

So, the metrics used in this case, when $L_k = H_k$, are

$$J_1^H = \frac{1}{N} \sum_{k=1}^N \text{tr} \{(A_{ak}^{11} + B_k^{\text{eff}} + R_k)^{-1} R_k\}, \quad (24)$$

$$J_2^H = \frac{1}{N} \sum_{k=1}^N \text{tr} \{(A_{ak}^{11} + B_{ak}^{11})^{-1} B_{ak}^{11}\} \quad (25)$$

with

$$A_{ak}^{11} = H_k F_{k-1} P_{k-1}^{H+} F_{k-1}^T H_k^T, \quad (26)$$

$$B_{ak}^{11} = H_k Q_{k-1} H_k^T; \quad B_k^{\text{eff}} = H_k Q_{k-1}^{\text{eff}} H_k^T. \quad (27)$$

An interesting observation from (27), which supports the earlier findings in [3], is that B_k^{eff} for the EHF can be obtained by replacing Q_{k-1} in B_k for the EKF by Q_{k-1}^{eff} as in (22).

More importantly, however, it is observed that since A_{ak}^{11} in (26) is dependent on P_{k-1}^{H+} , hence, J_2^H is dependent on the particular choices of γ and R_{k1} for the EHF and so it differs from J_2 as defined for the EKF. Furthermore, J_1^H in (24) also differs from the sensitivity metric J_1 as defined for the EKF since D_{ak}^{11} as in (23) has terms containing γ , R_{k1} , and P_{k-1}^{H+} in it. Hence, as desired, both J_1^H and J_2^H are dependent on the specific tuning parameters of the EHF and particularly on the robustness bound γ .

3 Problem Statement

The particular problem considered in this paper for evaluating the usable range of γ is that of tracking a 2D ballistic target, the system descriptions of which are the same as that used in [1]. In this paper, the desired states of the EHF are considered to be the same as the measurements in order to facilitate a comparison of the results of the EHF with those of the EKF as presented in [1].

3.1 Truth Generation

3.1.1 Process Model

The target engagement scenario, the process model description, and the system parameter values are the same as in Farina et al. [5]. It is assumed that the object enters the atmosphere in reentry phase under the presence of nonlinear air drag as well as gravity. The presence of drag as well as gravity makes the trajectory of the target a nonlinear one. The equivalent discrete-time target motion can be expressed [5] as

$$\begin{aligned} x_{k+1} &= f(x_k) + Gu_k + w_k \\ \text{where } f(x_k) &= F_1 x_k + Gf_{kk}(x_k). \end{aligned} \quad (28)$$

Here, x_k is the state vector consisting of the positions of the ballistic object in the x and y directions, denoted as x_{1k} and x_{2k} , respectively, and their corresponding velocities. Of the other terms, $f(x_k)$ is a nonlinear state transition relation and $u_k = [0 \ (-g)]^T$ is the input matrix with the associated structure matrix G , while $w_k = N(0, \sqrt{Q_t})$ is the process noise. w_k is assumed to be a zero mean white Gaussian noise with the fixed covariance matrix Q_t , which is dependent on T , the time interval between the radar measurements.

The nonlinear state transition $f(x_k)$ comprises of two parts as stated, and in that, F_1 is the linear part of the state transition matrix, while $f_{kk}(x_k) = -\frac{1}{2\beta} \rho g v \begin{bmatrix} \dot{x}_{1k} \\ \dot{x}_{2k} \end{bmatrix}$ is the nonlinear state-dependent function associated with the common structure matrix G . Here, β denotes the ballistic coefficient, ρ denotes the air density, and g denotes the acceleration due to gravity, while $v = \sqrt{\dot{x}_{1k}^2 + \dot{x}_{2k}^2}$ is the velocity of the ballistic object. The parameters x_k , F_1 , G , and Q_t are listed in Table 1.

The initial value of the state vector x_0 is $[232 \text{ km } 2,290 \cos(190^\circ) \text{ ms}^{-1} \ 88 \text{ km } 2,290 \sin(190^\circ) \text{ ms}^{-1}]^T$. The acceleration due to gravity g is assumed to be constant at 9.81 ms^{-2} . Hence, u_k becomes time-invariant, so it is henceforth referred to as the input matrix u . The ballistic target trajectory has been generated considering a constant target ballistic coefficient $\beta = 40,000 \text{ kgm}^{-1}\text{s}^{-2}$, while the time interval between the radar measurements is considered to be $T = 2 \text{ s}$. A crucial parameter in this problem is the air density function ρ which is considered to decay exponentially as $\rho = C_1 e^{-C_2 x_{2k}}$ with $C_1 = 1.227$, $C_2 = 1.093 \times 10^{-4}$, when $x_{2k} < 9,144 \text{ m}$, $C_1 = 1.754$, $C_2 = 1.490 \times 10^{-4}$, and when $x_{2k} \geq 9,144 \text{ m}$.

3.1.2 Measurement Model

The measurement equation is expressed as

$$y_k = h(x_k) + v_k \quad (29)$$

where $h(x_k)$ is the nonlinear function of the states and the measurement noise $v_k = N(0, \sqrt{R_t})$ is assumed to be a zero mean white Gaussian noise with its corresponding covariance matrix R_t .

Table 1 Process parameters for truth generation

x_k	F_1	G	Q_t
$\begin{bmatrix} x_{1k} \\ \dot{x}_{1k} \\ x_{2k} \\ \dot{x}_{2k} \end{bmatrix}$	$\begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} \frac{T^2}{2} & 0 \\ T & 0 \\ 0 & \frac{T^2}{2} \\ 0 & T \end{bmatrix}$	$\begin{bmatrix} \frac{T^3}{3} & \frac{T^2}{2} & 0 & 0 \\ \frac{T^2}{2} & T & 0 & 0 \\ 0 & 0 & \frac{T^3}{3} & \frac{T^2}{2} \\ 0 & 0 & \frac{T^2}{2} & T \end{bmatrix}$

The set of instantaneous measurements comprising of the range rate \dot{r}_{mk} , the elevation angle θ_{mk} , and the sightline rate $\dot{\theta}_{mk}$ have been generated by adding a white Gaussian noise $N(0, \sqrt{R_t})$ to the theoretical measurements $[\dot{r}_k \quad \theta_k \quad \dot{\theta}_k]^T$, which are obtained as:

$$\begin{bmatrix} \dot{r}_k \\ \theta_k \\ \dot{\theta}_k \end{bmatrix} = \begin{bmatrix} (x_{1k}\dot{x}_{1k} + x_{2k}\dot{x}_{2k}) / (\sqrt{x_{1k}^2 + x_{2k}^2}) \\ \tan^{-1}(x_{2k}/x_{1k}) \\ (x_{1k}\dot{x}_{2k} - x_{2k}\dot{x}_{1k}) / (x_{1k}^2 + x_{2k}^2) \end{bmatrix} \quad (30)$$

where

$$R_t = \begin{bmatrix} (15 \text{ ms}^{-1})^2 & 0 & 0 \\ 0 & \left(\frac{0.1^\circ}{57.3}\right)^2 & 0 \\ 0 & 0 & \left(\frac{0.01^\circ}{57.3} \text{ s}^{-1}\right)^2 \end{bmatrix}$$

4 Results and Discussion

As stated at the outset, the problem of tracking a 2D falling body as stated in [1] for the EKF is considered in this paper for evaluating the usable range of γ in the EHF. The truth for the states and the measurements have been generated using the same method as described in detail in [1]. The filter tuning parameters x_0^+ , P_0 , and Q_{nom} used for the EHF are also the same as the initial estimation error, its covariance, and the nominal process noise covariance as determined for the EKF in [1]. R_k is chosen to be time-invariant and same as that of the truth. An additional parameter to be chosen for the EHF is the desired output weighting matrix R_{k1} . Interestingly, this can be chosen arbitrarily since it appears as a product with γ in all cases. In this case, it is chosen to be identical to R_k without any loss of generality.

Using these, the effectiveness of the performance metrics, namely (J_1^H, J_2^H) , in predicting the natures of the corresponding RMSE performances has been verified in this paper for some particular combinations of Q_0 and γ . Furthermore, the regions of robust, sensitive, or balanced RMSE performances have been identified in a manner similar to that for the EKF from the metric plots for various choices of γ . The interpretation of the results ensure the predictability of the usable range of γ using the off-line metric plots.

4.1 Varying the Filter Parameters Q_0 and γ

The initial process noise weighting matrix Q_0 is varied in a manner similar to that used for the EKF by choosing different integer values of p in the expression $Q_0 = 10^p(Q_{\text{nom}})$. Accordingly, the value of n_q also changes. The typical filter tuning parameter for the EHF is γ . For the present problem, different initial values of

γ are considered for verifying its effect on these filters. These are denoted as $\gamma_1^i = c_i \times \|P_0\|$ where i represents the different choices of the initial values and c_i is an arbitrarily chosen constant value. Thereafter, the value of γ is changed at every k th instant using the relation $\gamma_k^i = c_i \times \|P_k^{H+}\|$, where P_k^{H+} is obtained from (11). The first choice of initial γ is considered to be $\gamma_1^1 = \infty$ in order to compare the corresponding RMSE performances with those of the EKF under similar conditions. Other values of γ_1^i have also been considered, and the specific choices of the constants c_i are 10^6 , 0.5, and 0.3 for $i = 2, 3$, and 4, respectively.

4.2 Simulation Results

The relevant sensitivity and robustness metrics of the EHF, namely J_1^H and J_2^H , have been calculated offline and plotted with respect to n_q for each $\gamma_1^i, i = [1, 4]$, as shown in Fig. 1. The corresponding RMSE performances of the EHF for all choices of γ are shown in Figs. 2, 3, and 4. The RMSEs are plotted for specific values of $n_q = 10.82, 3.82, 2.82, 1.82, 0.82$ and (-8.18) , and these are marked in the metric plots as $R3, R2, R1, S1, S2$, and $S3$, respectively. It is to be noted that the RMSEs for position y are not shown since they are similar in nature to those of position x in all cases.

Correlating the RMSE performances of the EHF for all choices of γ as shown in Figs. 2, 3, and 4 with the characteristics expected from the metrics in Fig. 1, it is observed that

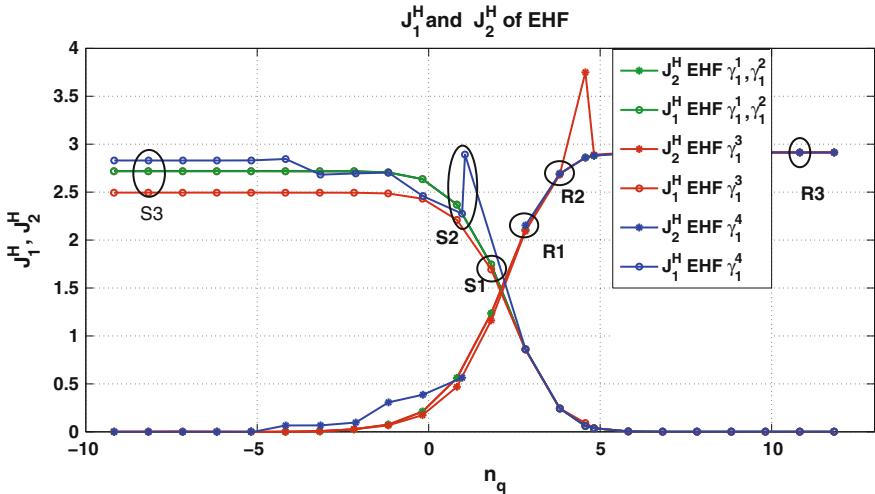


Fig. 1 Plots of J_1^H and J_2^H for different values of γ

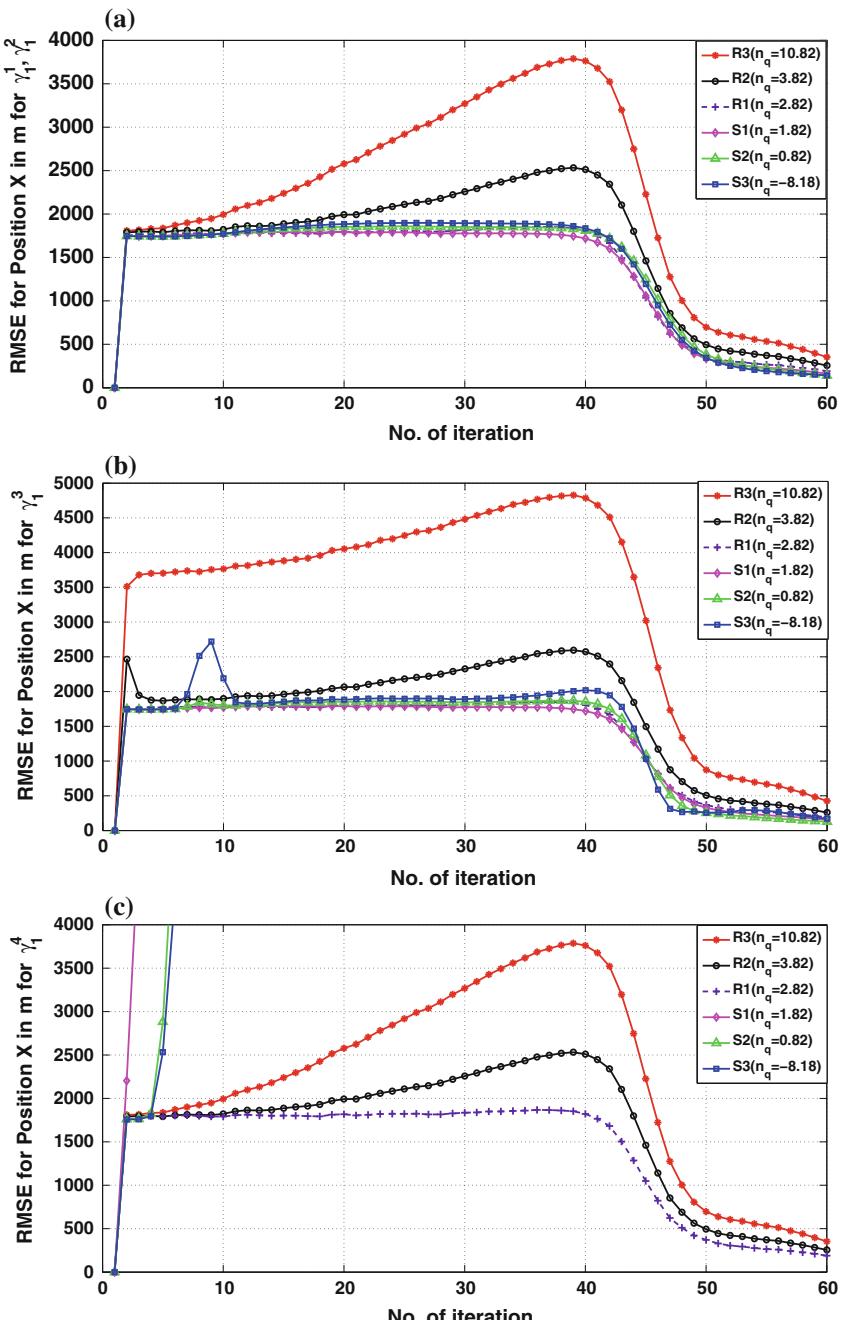


Fig. 2 RMSE plots of position x using EHF for various γ_1^i : **a** γ_1^1 , γ_1^2 , **b** γ_1^3 , and **c** γ_1^4

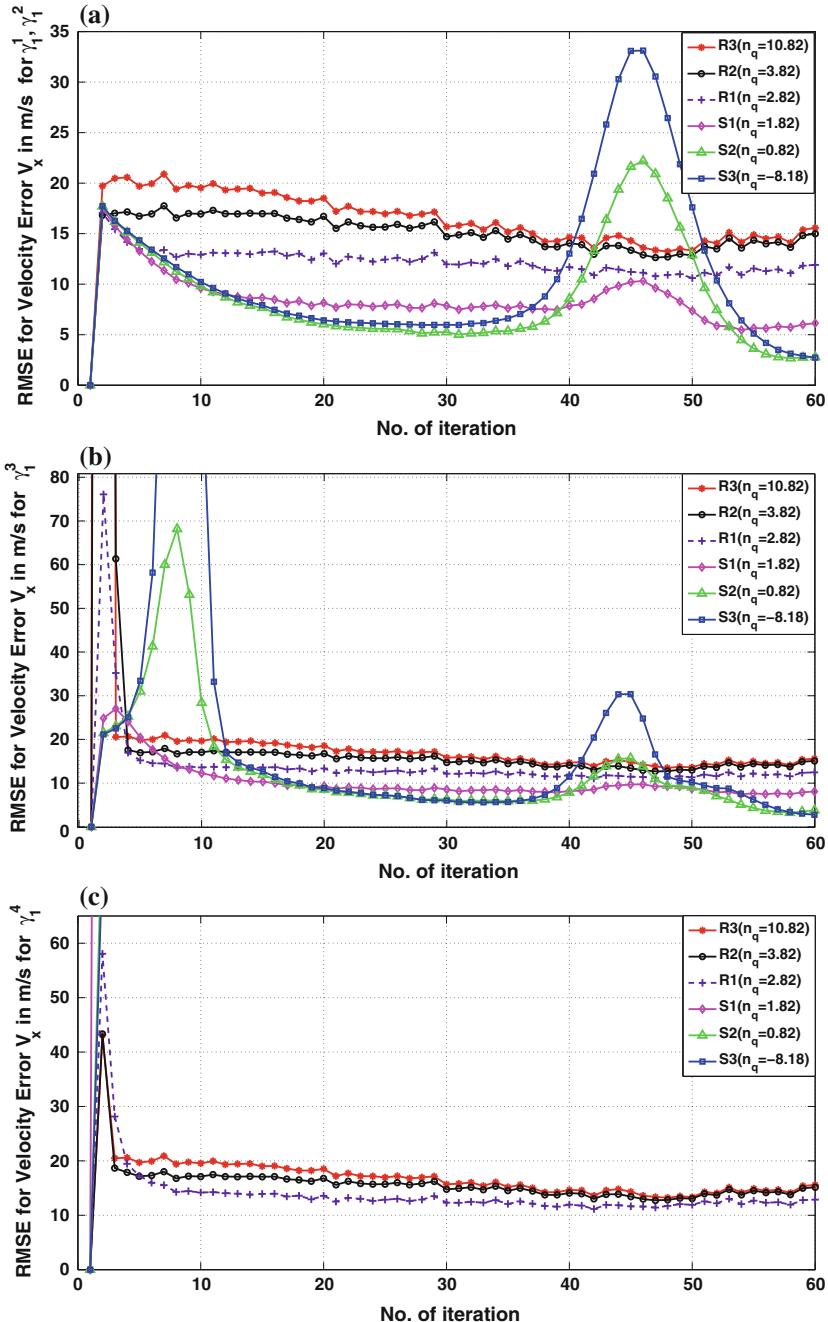


Fig. 3 RMSE plots of velocity v_x using EHF for various γ_1^i : **a** γ_1^1 , γ_1^2 , **b** γ_1^3 , and **c** γ_1^4

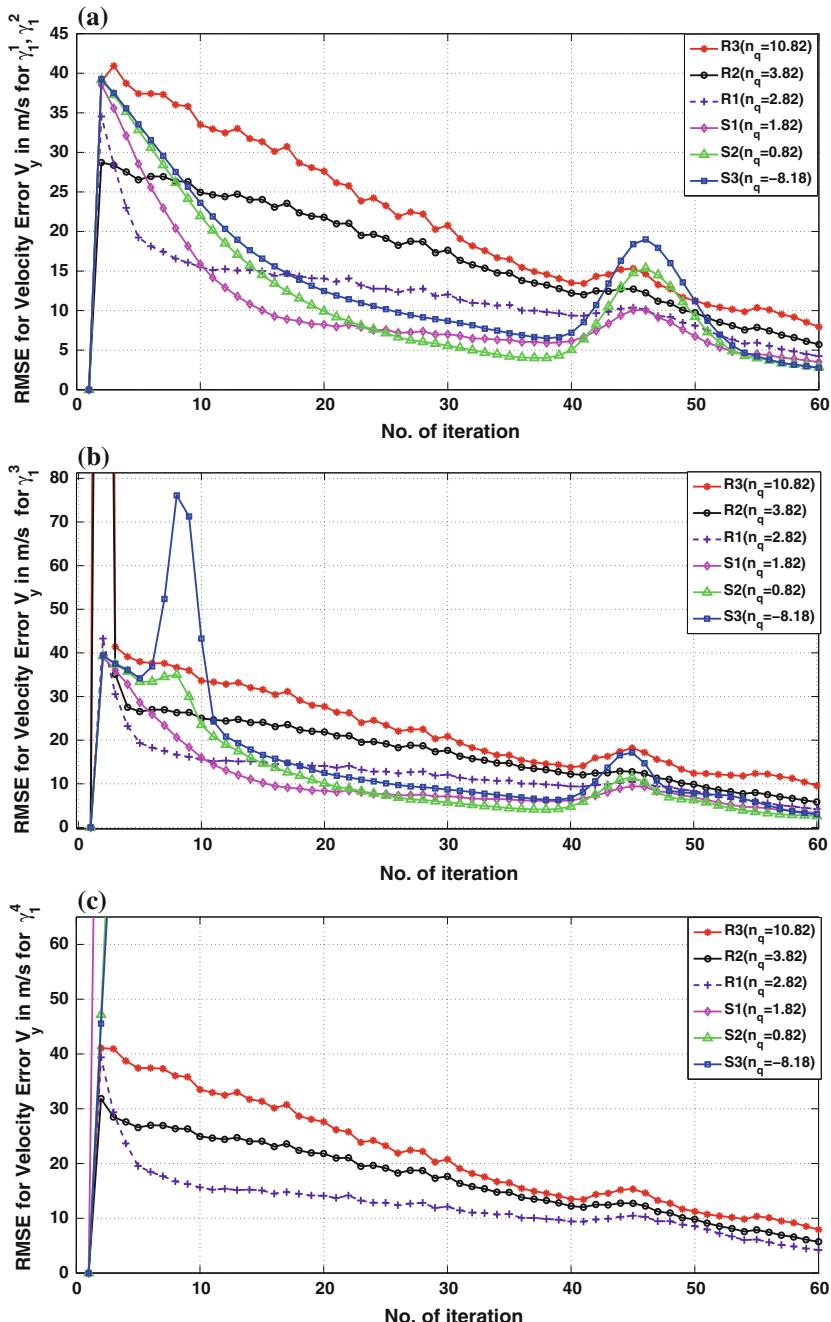


Fig. 4 RMSE plots of velocity v_y using EHF for various γ_1^i : **a** γ_1^1 , γ_1^2 , **b** γ_1^3 , and **c** γ_1^4

1. The robustness metrics J_1^H and J_2^H of EHF for $\gamma_1^1 = \infty$ are identical to J_1 and J_2 of EKF. Furthermore, since $\gamma_1^2 \rightarrow \infty$, in this case also J_1^H and J_2^H are identical to J_1 and J_2 of EKF.

Accordingly, as expected, the RMSE responses for $\gamma_1^1 = \infty$ and γ_1^2 are identical in all respects, as has been stated in available literature [2, 3]. Consequently, as in the EKF, in the EHF also, a high-valued bias-type error is observed in both the position estimates which can be ascribed to the lack of the range measurement. Also, the nonlinearity introduced at the altitude of 9,144 m seems to provide a corrective effect to this bias, and hence, the error reduces thereafter to acceptably low values.

2. For all these choices of γ , the metric plots for J_1^H (J_2^H) attain steady maximum and minimum values denoted as $J_{1\max}^H$ ($J_{2\max}^H$) and $J_{1\min}^H$ ($J_{2\min}^H$), respectively. For all these γ , $J_{1\max}^H$ is attained for $n_q \leq (-3.18)$ and this value decreases from 2.72 to 2.50, as the value of γ is decreased from γ_1^1 to γ_1^4 . Also, as expected, $J_{2\min}^H = 0$ for $n_q \leq (-4.18)$. On the other hand, $J_{2\max}^H = 2.91$ and $J_{1\min}^H = 0$ for $n_q \geq 5.82$ in all these cases. So, in the EHF also, there is a crossover of the two plots of J_1^H versus n_q and J_2^H versus. n_q for some particular value of the weighting matrix Q_0 , which can thus be identified as Q_{comp} for the filter.

Accordingly, in all cases, it has been independently verified that the filter performances are identically robust when $J_2^H = J_{2\max}^H$ and identically sensitive when $J_1^H = J_{1\max}^H$. Hence, the RMSE plots for $R3$ and $S3$ are the robust and sensitive performances, respectively, while the balanced performances are expected for $R2$, $R1$, $S1$, and $S2$. So, in all estimates, the robustness in the RMSE performances increases in comparison from $R1$ to $R3$ with $R3$ denoting the robust performance with a similar result holding true from $S1$ to $S3$ for sensitivity.

3. Lowering the value of γ substantially affects both J_1^H and J_2^H adversely in the form of a lack of monotonicity and/or boundedness in the plots at one or multiple values of n_q . This may be interpreted as an unreasonable increase in the robustness demand by the designer, which the system may not be capable of fulfilling. This adverse effect can be related to a corresponding increase in the term $(-\frac{1}{\gamma} R_{k1}^{-1})$, which tends to destabilize the filter by reducing the positive definiteness of M_{ak} in (8) and hence of P_{k-1}^{H+} as in (11).

The associated adverse consequences are observed in the RMSE performances of *all* the estimates pertaining to the sensitive zone for γ_1^3 , and particularly for γ_1^4 .

4. However, the effect of γ_1^i in enhancing the robustness of the EHF is prominent in the sensitive zone. This can be inferred from the observation that except for γ_1^4 , the values of J_1^H and J_2^H for the choices $S1$, $S2$, and $S3$ decrease monotonically as the value of γ_1^i is reduced, indicating reduced sensitivity in the performances. Accordingly, from the RMSE performances at $S1$, $S2$, and $S3$ for γ_1^1 till γ_1^3 , it is observed that as the value of γ_1^i decreases, the responses are similar but progressively less sensitive even to the nonlinearity introduced at the altitude of 9,144 m.

Correlating these observations from the metrics and the RMSE performances, it is possible to identify that for this problem, the permissible range of choice of initial values of γ is $[\gamma_1^1, \gamma_1^3]$, although the overall performance, from robust to balanced to sensitive, is observed to be the best for γ_1^1 , or in other words the EKF.

5 Conclusions

In this paper, the EHF has been studied as a modified EKF. The relevant robustness and sensitivity metrics for this class of filters have then been derived and evaluated in order to predict the usable range of robustness bound γ using the metrics which ensure acceptable filter performances. The particular application considered to test this is that of the realistic problem of tracking of a 2D falling body as discussed in [1].

For this, the modified EKF algorithm for the EHF as stated in [2] has been altered for applicability to suit any general problem description. Based on this modified algorithm, the augmented sensitivity and robustness metrics, namely J_{1a} and J_{2a} , have been derived for the general EHF analogous to those for the EKF. Thereafter, since the specific choice of the desired output vector z_k considered in this paper is that it is the same as the measured outputs, or $z_k = y_k$ and $L_k = H_k$, the corresponding set of sensitivity and robustness metrics, namely J_1^H and J_2^H , have been derived since J_{2a} becomes singular for this case.

Thereafter, the predictability of the RMSE performances of the EHF using J_1^H and J_2^H has been verified for different combinations of the tuning parameters Q_0 and γ . In the case of the EHF also, it has been established that a suitable combination of the relevant sensitivity and robustness metrics can be used to identify the regions of robust, sensitive, or balanced performances. Furthermore, the effect of changing the robustness bound γ from infinity to a low value, as interpreted in terms of the change in the relevant metrics, has been correlated to relevant changes in the RMSE performances, and accordingly, a prediction of the permissible range of γ for ensuring reliable filter performances is possible based on the off-line metric plot.

References

1. Saha, M., Ghosh, R., Goswami, B.: Robustness and sensitivity metrics for tuning the extended Kalman filter. *IEEE Trans. Instrum. Meas.* **63**(4), 964–971 (2014)
2. Dan, S.: Optimal state estimation: Kalman, H_∞ and nonlinear approaches. Wiley Interscience (2006)
3. Forssell, U.: On H_2 and H_∞ optimal estimation. <ftp://ftp.control.isy.liu.se/pub/Reports/1996/1885.ps.Z> (1996)
4. Petersen, K.B., Pedersen, M.S.: The matrix cookbook. Technical University of Denmark. <http://matrixcookbook.com> (2008)
5. Farina, A., Ristic, B., Benvenuti, D.: Tracking a ballistic target: comparison of several filters. *IEEE Trans. Aerosp. Electron. Syst.* **38**, 854–867 (2002)

Performance Evaluation of Forward Error Correction Schemes in Wireless Body Area Networks

K.S. Deepak and A.V. Babu

Abstract This paper considers the use of Luby transform (LT) codes as a forward error correction (FEC) scheme for wireless body area networks (WBANs). We evaluate the bit error rate (BER) and packet error rate (PER) performance of LT codes for the following communication scenarios in WBAN: (i) in-body communication between invasive devices and hub and (ii) on-body communication among noninvasive devices and hub with line of sight (LOS) and non-LOS (NLOS) channels. A WBAN in which sensor nodes monitor the physiological parameters of human body and sends the data to a common hub regularly is considered. We also evaluate the performance of Bose–Chaudhuri–Hocquengham (BCH) codes, the FEC scheme specified by IEEE 802.15.6 for WBANs.

Keywords BCH codes · IEEE 802.15.6 · Luby transform codes · Wireless body area networks

1 Introduction

Wireless body area networks (WBANs) consist of a number of low-power sensor nodes that are attached to the body surface, implanted inside or around the body, that monitor vital body parameters and transmit the information to a central device known as hub [1]. Medical applications that involve real-time collection of patient's vital information require almost 100 % reliability [2]. It has been reported that the wireless link between the sensor and the hub experiences severe attenuation of the order of 50–100 dB. WBANs experience deep fading effects that can last much

K.S. Deepak (✉) · A.V. Babu

Department of Electronics and Communication Engineering,
National Institute of Technology Calicut, Calicut 673601, Kerala, India
e-mail: p110043ec@nitc.ac.in

A.V. Babu
e-mail: babu@nitc.ac.in

longer (10–300 ms) than that experienced in cellular networks and severe shadowing effects that can result in very high outage probability [3].

The focus of this paper is to investigate bit error rate (BER) and packet error rate (PER) performance of forward error correction (FEC) schemes in WBANS. The IEEE 802.15 Task Group 6 has recently approved the PHY and MAC specifications for WBANs [4]. It defines a MAC layer in support of three PHY layers: Narrow-band (NB), Ultra wideband (UWB), and Human Body Communications (HBC). The standard specifies the use of BCH code ($n = 63, k = 51$) as the FEC scheme for the NB PHY. This coding scheme has capability to correct up to two bits of error per code word. In this paper, we evaluate the error rate performance when BCH codes are employed, considering in-body as well as on-body (LOS and non-LOS) channel models. Being a fixed rate coding scheme, use of BCH codes would result in decoding failures when channel conditions vary dynamically. Luby transform codes (LT codes) are the first class of practical rate less fountain codes that are near-optimal erasure correcting codes [5]. These codes have the property that any number of code symbols can be generated on the fly, and the data can be recovered from any set of code symbols slightly bigger than the data set, with high probability. In this paper, we consider the use of LT codes as a FEC scheme for WBAN and evaluate the error rate performance in realistic channel models applicable for WBANs.

Several papers have appeared in the literature that studies the reliability of WBANs [6–13]. In [6], outage performance and energy efficiency of direct transmission and single- and multi-relay cooperation schemes are analyzed in the context of WBANs. Energy efficiency as well as optimal packet size is analyzed in [7] for cooperative WBANs based on a realistic channel model. The energy efficiency of IEEE 802.15.6 compliant WBAN has been analyzed in [8] for the scheduled access mode. In [9], authors propose cooperative network coding to improve reliability of WBANs. Authors of [10] propose a new type of phase shift keying modulation scheme, namely phase silence shift keying (PSSK), for high data rate implantable medical devices and show that the BER of an 8-PSSK is lower than that of various sinusoidal carrier-based modulation schemes. However, in all the above-mentioned papers, FEC schemes are not considered for the analysis and evaluation. In [11], authors present a digital fountain code as a design criterion in order to mitigate packet erasure in WBANs. In [12], authors present an augmentation protocol for the physical layer of the MICS standard with focus on the energy efficiency and the reliability of LT coded NC-MFSK scheme over the MICS frequency band. In [13], authors calculate PER and propose a packet error reduction technique based on measured received signal strength for reliable communication in WBAN. Different from all these papers, our focus in the present paper is on the error performance of FEC schemes in WBAN. The rest of the paper is organized as follows. Section 2 describes the system model. Section 3 describes the results of the performance evaluation. Finally, Sect. 4 concludes the paper.

2 System Model

This section gives a brief description of LT codes and IEEE 802.15.6 NB PHY. The channel model used for the evaluation is also described. We consider a WBAN that is used for healthcare application where implant or body surface sensor nodes are used to monitor vital physiological parameters of a person and transmit the data to a central device known as hub located on the body surface.

2.1 LT Encoding and Decoding

The LT encoder is capable of generating an infinite stream of encoded symbols on the fly from k input symbols. The encoded symbol is generated using three steps: (i) determine the degree d (value between 1 and k) of encoded symbol from a degree distribution $\mu(d)$ based on the robust soliton distribution given in [5]; (ii) select d distinct input symbols uniformly at random as neighbors of the encoded symbol; and (iii) perform XOR of these symbols to form the encoded symbol. For decoding, it is assumed that the decoder has knowledge of the degree and neighbors of each encoded symbol. The decoding algorithm consists of following three steps [5]: (i) find encoded symbol connected to only one input symbol and the input symbol is recovered (decoding will fail if no such symbol exists); (ii) the recovered symbol is XORed with all encoded symbols that have this input symbol as neighbor; (iii) the recovered symbol is removed as neighbor from all these encoded symbols and their degree is reduced by one. Repeat above steps until all the input symbols are recovered.

The degree distribution $\mu(d)$ is a critical parameter of the fountain codes. Ideally, the decoding process should run in such a way that in every iteration, only one code word with degree one appears after an input symbol has been recovered. This can be achieved by the ideal soliton distribution given as follows [5]:

$$\begin{aligned} \rho(d) &= 1 && \text{for } d = 1 \\ &= 1/d(d-1) && \text{for } d = 2, 3, \dots, k \end{aligned} \quad (1)$$

This algorithm is designed such that it expects one encoded symbol to be recovered initially and exactly one input symbol is recovered in each iteration. However, this distribution works poorly in practice since any variation in ideal behavior is likely to cause degree-one encoded symbol to disappear and the decoding process to fail [5]. Luby proposed the robust soliton distribution which has two additional parameters c and δ to overcome this shortcoming. This distribution is obtained by adding a positive function τ given below to the ideal soliton distribution:

$$\begin{aligned} \tau(d) &= S/dk && \text{for } d = 1, 2, \dots, (k/S) - 1 \\ &= S/k \ln(S/\delta) && \text{for } d = k/S \\ &= 0 && \text{for } d = (k/S) + 1, \dots, k \end{aligned} \quad (2)$$

Here, S is the expected number of degree-one symbols given by Ref. [5]:

$$S = c \ln(k/\delta)/\sqrt{k} \quad (3)$$

for some suitable constant $c > 0$, and δ is the allowable failure probability. Notice that $\tau(d)$ is added to the ideal soliton distribution to get the robust soliton distribution $\mu(d)$ given as follows:

$$\mu(d) = (\rho(d) + \tau(d))/Z \quad (4)$$

where Z is a normalizing factor given by $Z = \sum_{d=0}^k (\rho(d) + \tau(d))$

2.2 IEEE 802.15.6 Physical Layer

The physical layer protocol data unit (PPDU) of NB PHY contains physical layer convergence protocol (PLCP) preamble, PLCP header, and a PHY service data unit (PSDU). The PSDU consists of MAC header, MAC frame body, and frame check sequence (FCS). The PSDU is encoded using binary BCH code and spread/interleaved before transmission. For an IEEE 802.15.6 compliant node, the standard specifies mandatory support of data rates with spreading and interleaving as given in Table 1 [4]. The standard specifies the use of BCH code ($n = 63$, $k = 51$, $t = 2$). The focus of this paper is on NB PHY in 402–405 MHz, 2.36–2.4 GHz, and 2.4–2.4835 GHz frequency bands. Medical implant communication takes place at 402–405 MHz band, and on-body communication takes place at 2.36–2.4 GHz and 2.4–2.4835 GHz bands [4].

Table 1 Modulation parameters for PLCP header and PSDU [4]

Frequency bands (MHz)	Packet component	Modulation	Symbol rate (ksps)	Code rate (k/n)	Spreading factor (S)	Information rate (kbps)	Support
2360–2400 and 2400–2483.5 (on-body)	PLCP header	π/2-DBPSK	600	19/31	4	91.9	Mandatory
	PSDU	π/2-DBPSK	600	51/63	4	121.4	Mandatory
	PSDU	π/2-DBPSK	600	51/63	2	242.9	Mandatory
	PSDU	π/2-DBPSK	600	51/63	1	485.7	Mandatory
	PSDU	π/4-DQPSK	600	51/63	1	971.4	Mandatory
402–405 (in-body)	PLCP header	π/2-DBPSK	187.5	19/31	2	57.5	Mandatory
	PSDU	π/2-DBPSK	187.5	51/63	2	75.9	Mandatory
	PSDU	π/2-DBPSK	187.5	51/63	1	151.8	Mandatory
	PSDU	π/4-DQPSK	187.5	51/63	1	303.6	Mandatory
	PSDU	π/8-D8PSK	187.5	51/63	1	455.4	Mandatory

2.3 Channel Model

The radio propagation in a WBAN takes place inside or in the vicinity of the human body. Channel modeling is challenging due to the complex shape and composition of the human body consisting of different tissues having different permittivity, conductivity, and characteristic impedance. We consider two distinct communication scenarios: (i) in-body communication between implant sensor nodes (nodes placed inside the human body) and hub and (ii) on-body communication between noninvasive sensor nodes (nodes placed on the human body) and hub. The non-invasive sensor node can have either line of sight (LOS) or non-LOS (NLOS) communication with the hub [14]. Since physical measurements and experimental studies inside the human body are extremely difficult, comprehensive characterization of in-body communication channel is tedious as compared to the on-body channel. The path loss models for in-body communication in the 402–405 MHz and on-body NLOS and LOS communication at 2.4 GHz are given by Refs. [14–16] as:

$$L_d = L_o + 10 \alpha \log(d/d_o) + X_\sigma \quad (5)$$

Here, d is the distance between transmitter and receiver; d_o is the reference distance; L_o is the path loss at reference distance d_o ; α represents the path loss exponent; and X_σ is a zero mean Gaussian random variable with standard deviation σ that represents the shadowing. The channel parameters for in-body, on-body NLOS and on-body LOS communication scenarios are obtained from [14–16], respectively, and are given in Table 2. The signal-to-noise ratio (SNR) at the receiver can be expressed as follows:

$$\gamma(\text{dB}) = P_t(\text{dB}) - L_d(\text{dB}) - P_n(\text{dB}) \quad (6)$$

where P_t is the transmit power and P_n is the additive white Gaussian noise power. It may be noted that the SNR expression given by Eq. (6) includes the effect of large-scale fading (represented by shadow fading random variable X_σ) as well. When LT

Table 2 The channel parameters for in-body, on-body NLOS and on-body LOS communication scenarios

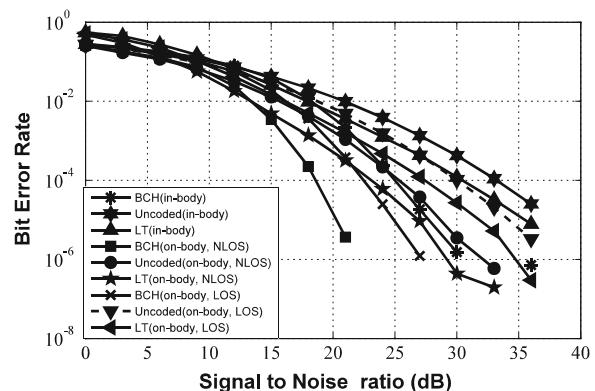
Transmit power	In-body	−10 dBm	
	On-body	−12 dBm	
Channel model parameters	Scenario	In-body	On-body NLOS
	d_o (cm)	5	10
	L_o (dB)	49.81	48.4
	α	4.22	5.9
	X_σ (dB)	6.81	5
Bandwidth	In-body	300 kHz	
	On-body	1 MHz	

coding is employed, the WBAN channel is assumed to be a binary erasure channel. It may be noted that the NB PHY layer specifications for a IEEE 802.15.6 compliant node specifies mandatory support of data rates with spreading and interleaving in addition to FEC. The coded bits are spread using a repetition code and then interleaved using a bit interleaver before modulation. As listed in Table 1, it can be observed that in the 2.4 GHz band, the spreading factor can be either 2 or 4 that leads to reduction of effective information data rate. It is assumed that, when spreading is done using repetition codes, majority logic-based detection is used at the receiver. Accordingly, the WBAN channel can be approximated as an erasure channel.

3 Performance Evaluation of FEC Schemes in WBAN

In this section, we evaluate the performance of FEC schemes (BCH and LT codes) in terms of BER and PER assuming realistic channel models for WBAN. Extensive Monte Carlo simulations are performed, assuming a WBAN with star topology in which data packets are transferred directly between the sensor node and the hub. Further, at the MAC layer, a contention-free MAC scheme such as scheduled uplink/downlink access mechanism [4] is assumed. The encoded bits are BPSK modulated and transmitted through the channel, where the symbols are affected by the distance-dependent path loss, noise, and shadowing effects corresponding to different propagation environments experienced in WBAN. We assume LT codes of rate comparable to that of BCH codes in which n^{LT} encoded symbols are generated from k input symbols using robust soliton distribution. The LT codes are rate less and n^{LT} can be potentially limitless. However, for simulation, we choose finite values for n^{LT} . This assumption is reasonable because, in practice, only finite symbols are necessary for decoding [5]. To have proper comparison against BCH codes, we take values of k and n^{LT} that are comparable to standard BCH code specifications [17]. Figure 1 shows the BER versus SNR when BCH coding

Fig. 1 BER versus SNR for BCH and LT coding schemes ($n = 63$, $k = 51$, $\varepsilon = 23.53\%$)



specified in the IEEE 802.15.6 standard is employed (i.e., $n = 63$, $k = 51$, $\varepsilon = 23.53\%$).

The results for uncoded and LT coded cases are also shown. Here, the distance separation between the nodes and the transmit power is fixed and the noise power is varied to obtain the desired SNR at the receiver. The corresponding specifications for LT code in terms of transmission overheads are defined as $\varepsilon = (n - k)/k$, i.e., code rate of $R_c = 0.8095$ is equivalent to overhead ratio of 23.53 %. Further, in-body communication has higher BER as compared to the on-body scenario, owing to the effects of absorption arising due to human body tissues, muscles, and fat which causes the path loss for the in-body channel to increase considerably. The BER performance of on-body NLOS communication is observed to be better than that of on-body LOS model due to higher value of shadow fading standard deviation observed in the on-body LOS propagation model. The performance of FEC schemes (BCH and LT codes) for different transmission overheads (40, 61.54, 98.44, and 195.35 %) was evaluated. Figure 2 shows the BER performance for $n = 127$, $k = 64$, and $\varepsilon = 98.44$ %. The BER is estimated for different values of code rate when the SNR is varied. In this case, BCH codes with different specifications, i.e., BCH (63,45,3), BCH (63,39,4), BCH (127,64,10), and BCH (127,43,14), were considered and their performance was compared against LT codes with the corresponding (n, k) values. It can be observed that, for lower overhead ratio, BCH codes outperform LT codes in terms of error performance. However, when the overhead ratio is increased, LT codes exhibit higher coding gain.

Figure 3 shows the coding gains achievable for a desired error rate of 10^{-5} when the transmission overhead is varied. Highest coding gain is observed for in-body propagation model, while lowest coding gain is observed for NLOS propagation model. This happens because in-body communication has the highest BER, while the BER of on-body NLOS case is the lowest.

Table 3 compares the SNR required at the receiver to achieve an error performance of 10^{-4} for BCH and LT codes at different code rates. It is observed that the SNR required decreases as the code rate decreases for both the coding schemes;

Fig. 2 BER versus SNR for BCH and LT coding schemes ($n = 127$, $k = 64$, $\varepsilon = 98.44\%$)

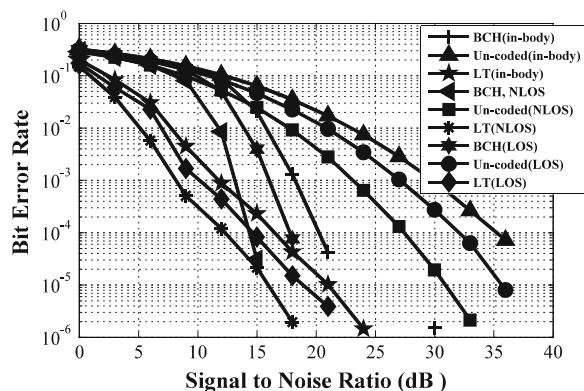


Fig. 3 Coding gain versus overhead for BCH and LT coding schemes ($\text{BER} = 10^{-5}$)

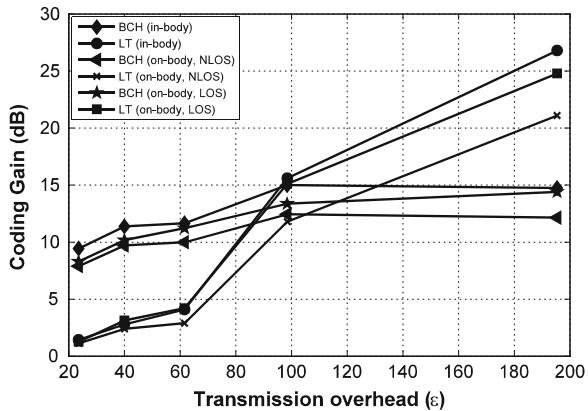


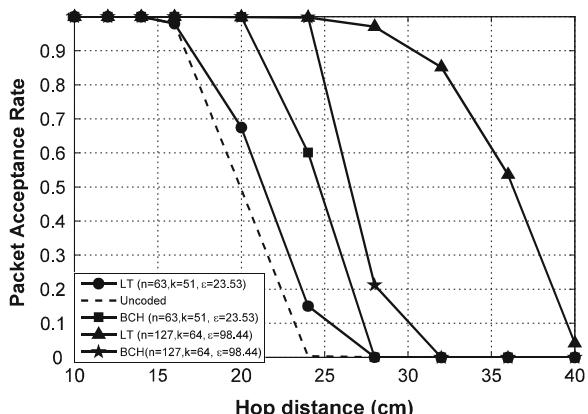
Table 3 SNR (dB) required for a desired error performance of 10^{-4} at different code rates

Transmission overhead ε (%)	In-body		On-body NLOS		On-body LOS	
	BCH	LT	BCH	LT	BCH	LT
23.53	24.90	30.50	18.90	22.94	22.30	27.50
40.00	23.35	28.45	17.50	21.61	21.00	26.21
61.54	23.02	27.00	16.10	20.80	19.80	24.70
98.44	20.08	16.33	14.70	12.00	18.00	13.90
195.35	19.35	05.49	14.70	03.50	17.26	04.34

however, this decrease is quite significant for LT codes. This shows that the transmit power required when LT codes are used significantly lower.

Figure 4 shows the reliability expressed as packet acceptance rate (PAR) for BCH and LT coded transmissions. The PAR decreases with hop distance owing to the higher path loss experienced. When the code rate is comparatively higher, BCH

Fig. 4 PAR versus hop distance for on-body NLOS channel



codes exhibit higher PAR as compared to LT codes. When the code rate is lowered, LT codes show better PAR performance than BCH codes. For the on-body NLOS model, to achieve a PAR equal to 0.9, use of BCH (127, 43, 14) extends the reliable hop distance by 7 cm, whereas LT codes with the same overhead ratio extend the hop distance by 13 cm.

4 Conclusion

In this paper, we have analyzed the error performance of WBANs when LT codes are used as FEC scheme. A WBAN in which sensor nodes send data to a common hub regularly was considered. For realistic channel models applicable for WBANs, it was verified that both BCH and LT codes can improve the performance of the network when compared to an uncoded system, owing to the coding gain they provide. However, it was observed that the performance improvement strongly depends on the amount of transmission overhead used. When overhead is increased, LT codes exhibit higher coding gain as compared to BCH codes, leading to significant reduction of the transmit power required for a given BER performance. However, the improvement in reliability is obtained at the cost of reduced bandwidth efficiency. This can be justified by the fact that, for WBANs, power constraint is much more critical than data rate due to the restricted battery lifetime of sensor nodes.

References

1. Cao, H., Leung, V., Chow, C., Chan, H.: Enabling technologies for wireless body area networks: a survey and outlook. *IEEE Commun. Mag.* **47**(12), 84–93 (2009)
2. Boulis, A., Smith, D., Miniutti, D., Libman, L., Tselishchev, Y.: Challenges in body area networks for healthcare: the MAC. *IEEE Commun. Mag.* **50**(5), 100–106 (2012)
3. Miniutti, D., Smith, D., Hanlen, L., Zhang, A., Boulis, A., Rodda, D., Gilbert, B.: Sleeping channel measurements for body area networks. *IEEE 802.15-09-0778-01-0006*, Jan 2010
4. IEEE Standard for Local and metropolitan area networks Part 15.6: Wireless Body Area Networks, 29 Feb 2012
5. Luby, M.: LT codes. In: Proceedings of IEEE Symposium on Foundations of Computer Science, pp. 271–280, Vancouver, Nov 2002
6. Huang, X., Shan, H., Shen, X.: On energy efficiency of cooperative communications in Wireless Body Area Networks. In: Proceedings of IEEE Conference on Wireless Communications and Networking Conference, pp. 1097–1101, Mexico (2011)
7. Deepak, K.S., Babu, A.V.: Packet size optimization for energy efficient cooperative Wireless Body Area Networks. In: Proceedings of IEEE India Conference (INDICON), pp. 736–741, Kochi, Dec 2012
8. Deepak, K.S., Babu, A.V.: Energy efficiency of IEEE 802.15.6 based wireless body area networks in scheduled access mode. In: Proceedings of Advances in Computing, Communications and Informatics (ICACCI), pp. 301–307, Mysore, Aug 2013

9. Arrobo, G.E., Gitlin, R.D.: Improving the reliability of wireless body area networks. In: Proceedings of IEEE Conference on Engineering in Medicine and Biology Society, pp. 2192–2195, Boston, 2011
10. Oh, J.-Y., Kim, J.-H., Lee, H.-S., Kim, J.-Y.: PSSK modulation scheme for high data rate implantable medical devices. *IEEE Trans. Inf. Technol. Biomed.* **14**(3), 634–640 (2010)
11. Takizawa, K., Hamaguchi, K.: Design of packet erasure mitigation technique using a digital fountain code for wearable wireless body area networks. In: Proceedings of IEEE Conference on Engineering in Medicine and Biology Society (EMBC), pp. 356–359, Buenos Aires (2010)
12. Abouei, J., Brown, J.D., Plataniotis, K.N., Pasupathy, S.: Energy efficiency and reliability in wireless biomedical implant systems. *IEEE Trans. Inf. Technol. Biomed.* **15**(3), 456–466 (2011)
13. Hamada, Y., Takizawa, K., Ikegami, T.: Highly reliable wireless body area network using error correcting codes. In: Proceedings of IEEE Radio and Wireless Symposium (RWS), pp. 231–234 (2012)
14. Yazdandoost, K.Y., Sayrafian-Pour, K.: Channel Model for Body Area Network (BAN). *IEEE P802.15-08-0780-09-0006* (2009)
15. Fort, A., Ryckaert, J., Dessel, C., De Doncker, P., Wambacq, P., Van Biesen, L.: Ultra-wideband channel model for communication around the human body. *IEEE J. Sel. Areas Commun.* **24**(4), 927–933 (2006)
16. Reusens, E., Joseph, W., Latré, B., Braem, B., Vermeeren, G., Tanghe, E., Martens, L., Moerman, I., Blondia, C.: Characterization of on-body communication channel and energy efficient topology design for wireless body area networks. *IEEE Trans. Inf. Technol. Biomed.* **13**(6), 933–945 (2009)
17. Lin, S., Costello, D.J.: Error Control Coding, 2nd edn. Prentice-hall, Englewood Cliffs (2004)

Simulating Self-Recovering Electric Circuits Using Neural Networks

Arijit Ray, S.S. Dash and N. Chellammal

Abstract Neural circuits and their applications for retrainable circuits have created new avenues in circuit design and fault correction. Electric systems and electronic circuits can experience a number of faults due to faulty connection of wires, misfiring of switches, wear-out of certain components and so on. Unlike a conventional circuit, a neural circuit is not affected by a single-component failure and thus can self-correct the fault. Once a fault occurs, the circuit can be retrained almost instantly to obtain the required output with the remaining healthy circuit once again. In fact, the fidelity is even better sometimes as overfitting is reduced. In this paper, a filter and an inverter were subjected to faults with more than half the transistors removed. The simulated circuits retrained in seconds to obtain the desired characteristics with almost the same accuracy. This paper delineates the various aspects of such a self-recovering electric system and the magnitude of the fault that can be recovered from.

1 Introduction

Artificial neural networks (ANNs) are widely known for their excellence in pattern recognition and modeling of complex relations [1]. This bio-inspired artificial intelligence tool has gained considerable popularity, both in terms of software and hardware research [2]. Electric systems can be modeled mathematically if the physics of each component is known. However, for increasingly complex systems, the mathematical modeling may be tedious and inaccurate due to various black box assumptions. Neural networks are excellent approximants and thus can model circuits very well as was shown in our previous paper [3]. Conventional circuits are highly dependent on each and every component of itself and thus have the hazard of failing due to a single-component failure. However, neural circuits are trainable

A. Ray (✉) · S.S. Dash · N. Chellammal

Department of Electrical and Electronics Engineering, SRM University, Chennai, India
e-mail: arijit.ray93@gmail.com

and reconfigurable and are thus not affected by a single-component failure. The circuit trains itself with the remaining healthy part to achieve the desired characteristics. Thus, a motor driver controlling several motors will not be affected even if a circuit component gets damaged. Such circuits can find immense applications in industrial drives and electronic systems where fault is intolerable, such as airplanes and race cars.

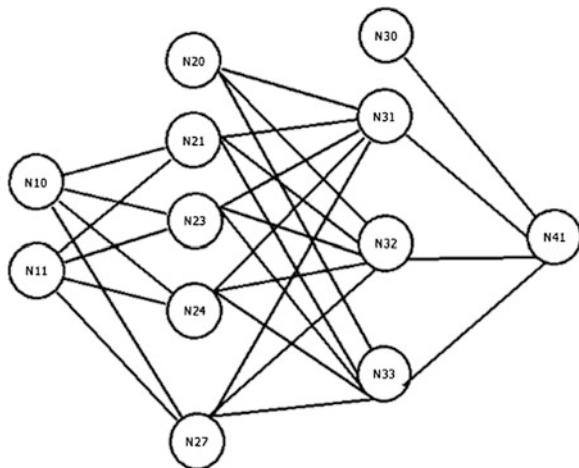
1.1 Related Work

Steven M. Bowers et al. at California Institute of Technology have attempted at creating self-healing amplifiers. They have shown that even after creating severe faults by destroying transistors from their circuit, the amplifier was able to retain its characteristics [4]. The circuit reroutes itself to retain the characteristics due to the redundancy present. However, in neural circuits, the output is reconfigurable due to the retrainable property of neural networks, and thus, it cannot exactly be called a redundancy. Neural circuits are also widely described and implemented in the literature. The fast and efficient circuit fabrication techniques nowadays make it possible for implementing circuits with millions of neurons in a few millimeters. The efficient reconfigurability of neural circuits is also widely discussed. Yun et al. described a reconfigurable and expandible architecture for implementing hardware neural networks. The number of layers can also be controlled using such architecture, and thus, such systems are useful for rerouting the circuit when fault occurs [5]. Recurrent neural networks such as Hopfield networks can model dynamic systems accurately. However, in this paper, we show an effective way of modeling time-dynamic circuits using simple feed-forward architectures only by using a time stamp input. Various genres of neural networks are also shown to be effectively implemented. John C. Gallagher et al. in their paper implemented a hardware form of continuous time recurrent networks. They demonstrated the applicability of such networks for trainable evolvable hardware [6]. However, usage of such circuits for creating self-healing circuits is little discussed. In this paper, we demonstrate an effective way of using simple feed-forward networks for modeling dynamic systems and provide a reconfigurable approach so that the circuit retrains itself after a fault occurs.

2 Proposed System

The neural circuit is shown to detect and recover from a deliberately caused fault and thus attain the pre-fault performance. The work is divided into the following parts:

Fig. 1 Neural network structure



- Training of a neural circuit to attain a particular response.
- Deliberately creating single- or multiple-component failure faults one by one.
- Detecting the location of the fault.
- Retraining the circuit and checking the retraining time to see how long the circuit will be in a faulty state.
- Checking the output of retrained circuit to check the error from the desired output.

3 Basic Construction and Training of a Neural Circuit

A neural network represents an interconnected network model consisting of nodes and branches (Fig. 1). The branches are associated with linear weights that are multiplied with the values output from the nodes. Each node has a sigmoidal transfer function. Such an interconnection makes it possible for the neural network to model highly complex relations and functions between input and output data. The input is given to the first layer, and the values are feed-forwarded to the reach the output value.

Feed forward is computed using the formulae:

$$a_2 = \text{sigmoid}(a_1 * \Theta_1) \quad (1)$$

Based on the required output, the cost is calculated using the cost function:

$$J = \left(\frac{1}{m}\right) \sum_{i=1}^m \sum_{k=1}^K (y - a_o)^2 + \frac{\lambda}{2m} \sum_{j=1}^{\text{theta size}} \text{internal thetas}^2 \quad (2)$$

The error is back-propagated to reach the optimal set of weights. For back-propagating, the derivative of the error function for each layer is calculated. For the output layer,

$$\delta = (a_o - y) \quad (3)$$

And for the subsequent layers,

$$\delta^l = (\theta^l)^T \delta^{l+1} * (g'(z^l)) \quad (4)$$

In hardware, a neural network can be easily modeled using field-effect transistors (FETs) and variable weight gain branches as described and explained in one of our previous papers [3]. FETs generally are known to have a sigmoidal transfer function with respect to the gate voltage and drain current. Thus, a FET can serve as a neuron node. Op-amps with electronically controllable gain are used as weighted branches for feed-forwarding values. Figure 2 represents the hardware form of the neural network with each branch as an op-amp gain and each node as a FET device. Each node has a node address with first digit as layer number and second digit as node number in the layer. Thus, N12 refers to the second node in the first layer. Similarly, a branch address can be written as a digit representing the origin layer succeeded by two numbers representing the origin node and the destination node. Thus, B234 represents the branch connecting the 3rd node of the 2nd layer to the 4th node of the 3rd layer. The optimized gain of each branch is determined and uploaded by electronically setting the gain of each branch. Such circuits can be used as self-constructing circuits that configure themselves based on the requirements.

The neuro-circuit can be trained in software by any matrix operation-enabled programming environment and then uploaded onto the circuit for the desired performance. Thus, using the method described in Fig. 3, the circuit in Fig. 2 can be trained for any circuit with the desired input and output characteristics. However, a simple feed-forward ANN can only model time-invariant and static waveforms. Thus, to model dynamicity, a time stamp input is added to the neural circuit in the form of a sawtooth wave.

The time period of the sawtooth wave should depend on the level of time variance required as is explained in one of our papers [3]. The ramp in the sawtooth wave can help in modeling time invariance in the output wave for that period of time. This is because the sawtooth time stamp creates a unique identifier for the particular input voltage to be modeled by the neural circuit to predict the required output. This is very much needed in time-variant circuits, as the output may be different for a same input at different times.

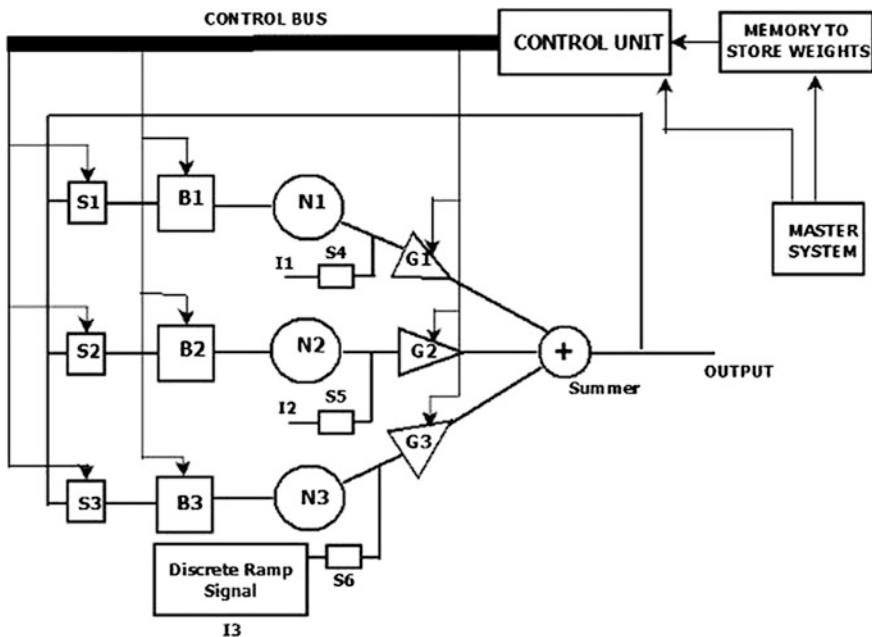


Fig. 2 Reconfigurable neuro-morphic circuit

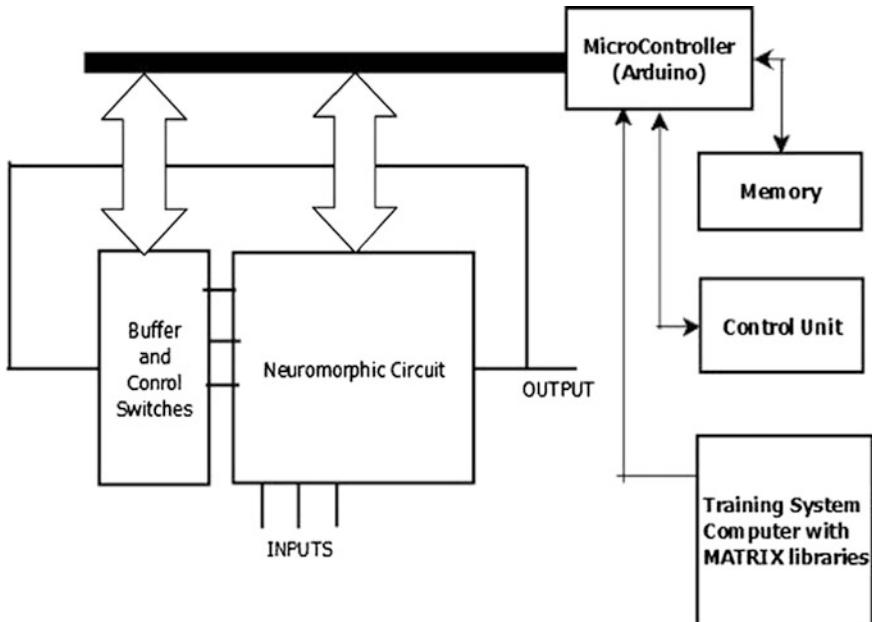


Fig. 3 Training and uploading process

3.1 Healthy Circuit Performance

In this paper, an inverter and a filter were modeled and the neural circuit shown in Fig. 2 was simulated in a MATLAB environment. The circuit was assumed to have no nonlinearities, and the transistors were assumed to have a perfect sigmoidal transfer function. However, in reality, there are errors induced in circuits and the transistor transfer functions are not accurately sigmoidal. This issue will be addressed in a future paper with a neural network to train live taking into account the various inadvertent errors in the circuit. The inverter is modeled using a direct voltage as input and a sinusoidal voltage as the output. The ramp signal input is matched to the time period of the sine-wave output required for maximum accuracy. Figure 4 shows an inverter modeled with the help of the neural circuit. The error of this healthy circuit is about 0.5 % in simulation. The fidelity of the waveforms is shown in Fig. 5. The output wave is shifted up for better understanding.

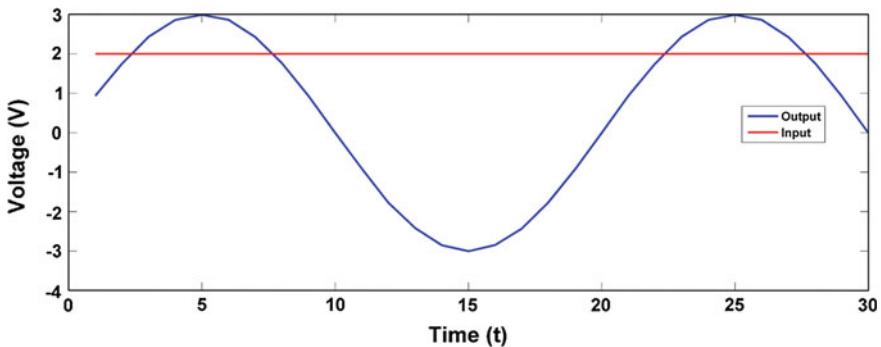


Fig. 4 Inverter modeling

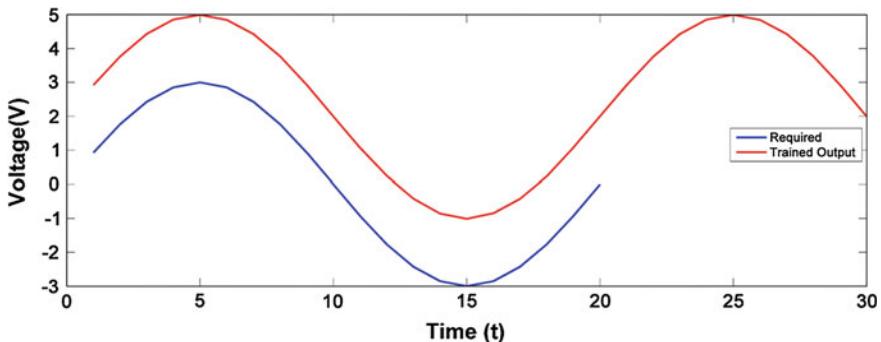


Fig. 5 Fidelity of inverter waveforms

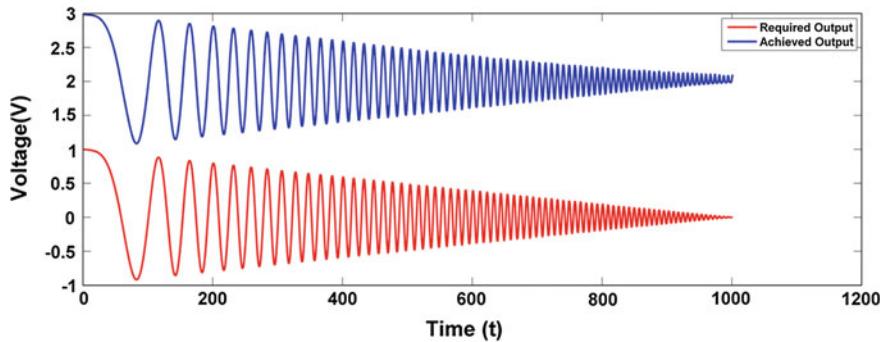


Fig. 6 Filter output

Similarly, a filter is modeled with a linear decay. The input is a waveform of increasing frequency commonly known as a chirp signal. The frequency of the input is varied from 1 to 150 kHz, and the required output is a linearly decaying signal to almost zero output for 150 kHz and beyond. The neural network is trained with the data, and it was successfully able to model it with an accuracy of 99.4 % in simulation as shown in Fig. 6.

4 Fault Recovery

Neural networks can model functions accurately even if their architecture is altered. For most general functions, a two-layered neural network with more than 30 neurons in each layer serves pretty well. Our circuit used for simulation used 40 neurons in each layer. Thus, if a fault occurs in the circuit, the remaining healthy circuit can be retrained to obtain the desired performance once again.

4.1 Types of Faults

There can be two types of faults in the neural circuit (Fig. 7):

- Nodal fault.
- Branch fault.

In a *nodal fault*, one or multiple nodes get dysfunctional. The FET might burn or get disconnected due to some reason causing a node loss to the circuit. This will cause wrong feed-forwarding of values and thus a fault in the output waveform. This is a benign fault compared to the branch fault and causes almost no harm at all to the circuit. Since a neural circuit can train itself very accurately for most waveforms with even 6–7 neurons in each layer, the circuit can sustain a high

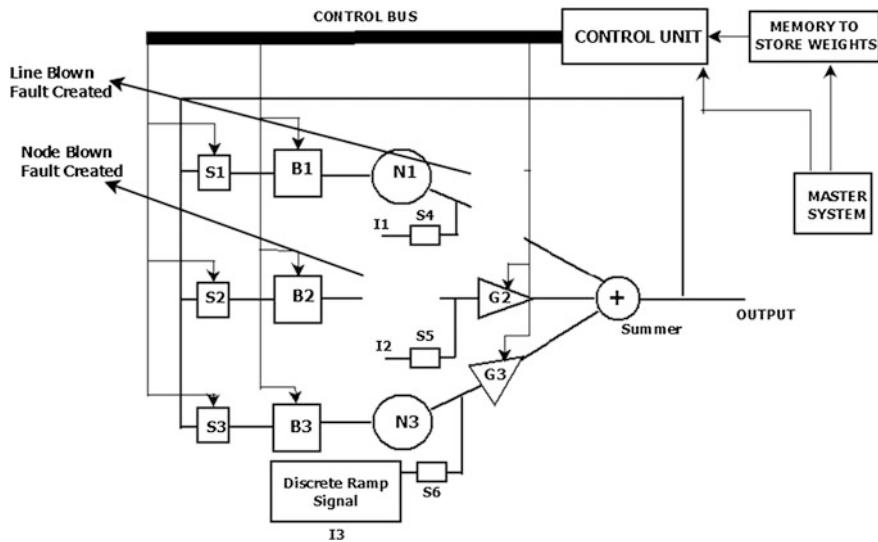


Fig. 7 Fault created

magnitude of this fault. The circuit here has 40 neurons in each layer and thus can sustain a fault of about 20 neurons.

A *branch fault* is where single or multiple branches get disconnected from a node, but the node remains healthy. This again causes a wrong feed-forwarding of values. It can even cause an error while back-propagating and thus deter it from converging. Thus, these faults can be more malignant and may have sustained recovery times. A naïve solution would be to cut off the entire node and retrain. However, if multiple branches are affected at multiple nodes, such a method can cut off a huge number of nodes unnecessarily and cause the circuit to fail. Thus, back-propagation has to be carried out by forcing the weights of the cut-off branches to be set to zero.

4.2 Error Detection

First, the fault is to be detected, which can easily be done by noticing the large deviation from the required output. An error threshold would be set. As soon as the output error increases the threshold, it means a fault has occurred in the circuit. However, the node/branch that has to be removed during training in software has to be detected. A node failure occurs due to the FET failure. Thus, there will be a high impedance in the source-drain junction in spite of a high gate signal. This can be utilized to detect the node that has failed. An FET monitor is fitted at every node which measures the source-drain current. As soon as the error comparator senses a

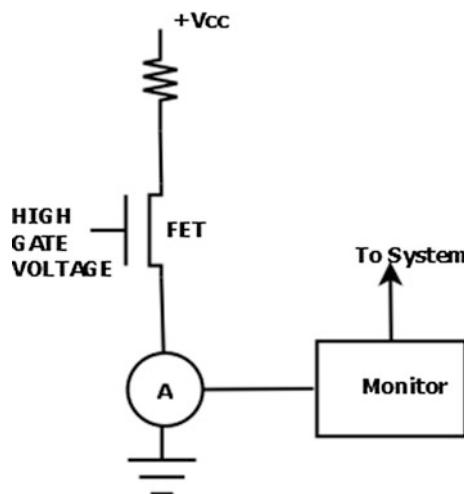
high deviation in the output, the input is first cut off. Then, all branch gains in the neural circuit are set to unity and all inputs are set to HIGH. This ensures a HIGH gate signal to all the neurons. Now, the source-drain current is monitored and the node exhibiting high impedance between drain and source is the faulty node. The node address is sent to the system as the faulty node address (Fig. 8).

Similarly, branch error can also be detected using a branch current detector. The faulty branch address is sent to the system and then it is decided whether to remove the entire node or train with branch gains set to zero. If all the branch failures are concentrated at a single or few nodes, then it is more profitable to remove the nodes and retrain. However, if the branch faults are distributed across various nodes, the circuit is retrained by forcibly setting the faulty branch gains as zero. The circuit is now retrained until the output error falls below a certain set threshold. As neural weights are randomly initialized at first, it is not always that the training will optimize to a global minimum; thus, training again and again is of paramount importance to achieve a perfect fidelity.

4.3 Post-Fault Performance

Post-fault performance is generally said to be lower than pre-fault performance for most conventional electric systems. However, in a neural circuit, the post-fault performance can in fact be better. This is because neural network training accuracy depends on various factors. Overfitting is a major cause for poor accuracy on test set. Complex neural networks sometimes tend to overtrain themselves and thus attain overfitted characteristics. Bridging a few branches or nodes reduces this problem, and thus, the circuit attains a more perfect fidelity. In this paper, a nodal

Fig. 8 Detecting nodal error



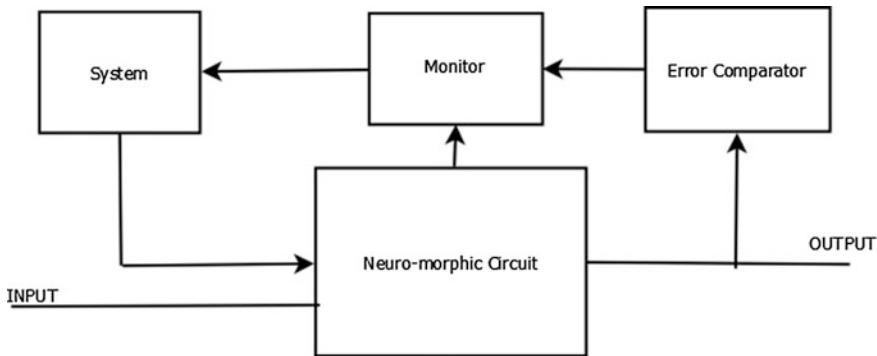


Fig. 9 Error recovery flowchart

fault and then a branch fault were created in both the inverter and filter and they were retrained. While the nodal fault caused no difference in the output, the branch failure sometimes caused multiple failures for back-propagation while converging. However, multiple tries with random weight initializations helped in retraining the circuit to regain the desired characteristics. If, after a threshold of retries, the training failed, the node was removed and the circuit was retrained for achieving the desired performance (Fig. 9).

4.4 Simulations of Post-Fault Performance

4.4.1 Inverter After Recovery from Fault

The healthy inverter circuit model had an error of 0.529 %. Nodal faults were created in the circuit. Even after damaging more than half of the nodal transistors, the circuit error just increased to 1 %. The fidelity was still undisturbed to plain sight as is shown in Fig. 10. The circuit showed a consistent performance even beyond the training set on a test set as is depicted by the actual curve in Fig. 10. The time required for the training algorithm to rerun in MATLAB was just 1.7 s, and it can go down to milliseconds in stand-alone neural circuit systems. This shows prospects for self-healing circuits that can withstand a lot of damage and repair itself almost instantly.

4.4.2 Filter After Recovery from Fault

The filter was also subjected to branch and nodal faults. In one case, the only three branches affected were removed, and in the other, the nodes were completely removed due to multiple branch failure. Surprisingly, the post-fault circuit fidelity

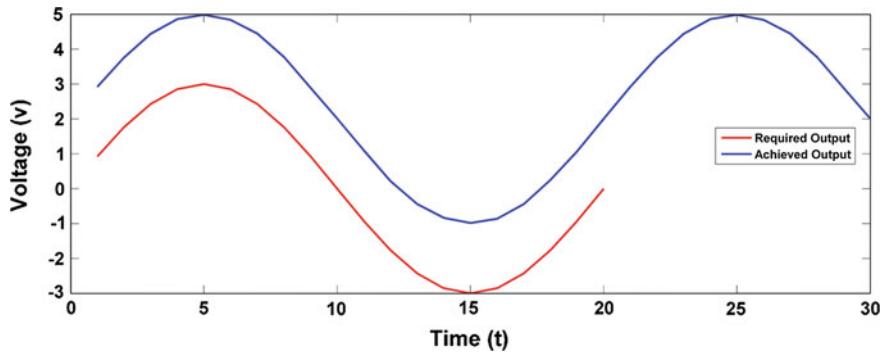


Fig. 10 Inverter after recovery from nodal error

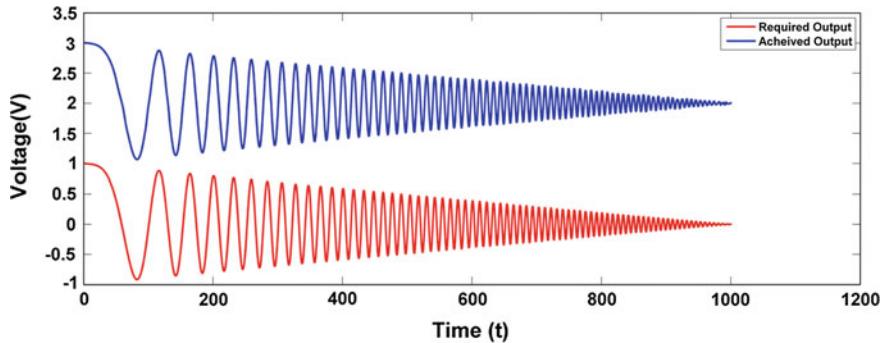


Fig. 11 Filter circuit after recovery from branch error

was sharpened as overfitting was reduced. After cutting off three branches, the error reduced to 0.5 % from the error of 0.6 % pre-fault. Figure 11 depicts the characteristic after recovery from branch fault. When multiple nodes were removed, the error increased slightly to 1.3 %. However, the fidelity of the waveform is pretty accurate to the naked eye.

4.4.3 Variation of Accuracy with the Magnitude of Fault

The variation of error with the number of transistors is shown in Fig. 12. It can be seen that even for low number of transistors the system error is pretty low. Some occasional spikes are seen for particular number of neurons, and it is because of its inability to converge to a global minimum.

However, this can be easily corrected by training again with another set of randomized weights or by removing another node. Thus, it quite implies from the graph that, after damage, not only can the circuit regain its previous performance,

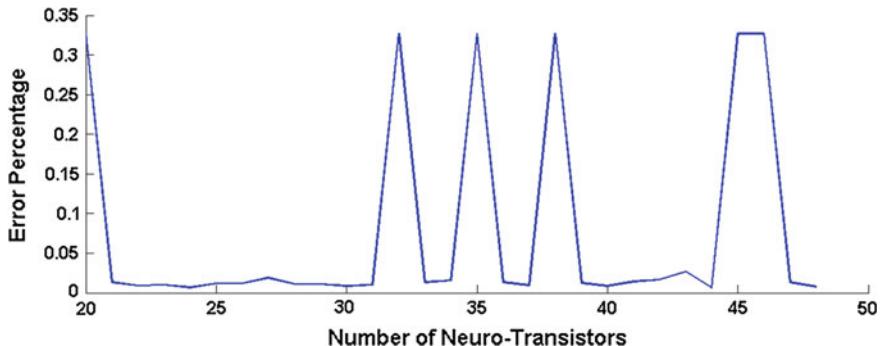


Fig. 12 Variation of error with the number of healthy neuro-transistors

but it also can even become more accurate. It can easily be inferred from the graph about the magnitude of the fault that can be recovered from. Even after more than half the neuro-transistors are destroyed from the original 40 that were present, the circuit mostly regains its full performance and in fact improves at times.

5 Conclusion

According to van Keulen et al. [7], a neural circuit must be sufficiently general and problem independent. This paper has proved the power of a neural circuit as a generalized as well as a self-healing one. An inverter and a filter were able to recover within seconds from a severe fault with more than half of the circuit components damaged. This opens avenues for constructing circuits that heal on the go. It can be used in electronic gadgets that self-repair themselves even if an internal short circuit occurs, thus obviating the need of servicing. This technology finds even more important applications in the field of avionics and industrial systems where a sustained fault can cause a huge loss. Large computer servers and electronic systems in airplanes can undergo self-healing in times of emergency and thus prevent system failure.

References

1. Le Cun, Y., Boser, B., Denker, J.S., Henderson, D.: Handwritten digit recognition with a back-propagation network. In: Proceedings of Advanced Technology Conference, AT&T Bell Laboratories, pp. 1003–1011 (1990)
2. Nawrocki, R.A., Shaheen, S.E., Voyles, R.M.: A neural architecture from single transistor neurons with organic bistable devices for weights neural networks (IJCNN). In: The 2011 International Joint Conference, pp. 450–456

3. Ray, A., Dash, S.S., Chellammal, N.: Reconfigurable time-variant electric circuits using time-stamped feed-forward neural networks, at IEEE Potentials (in review)
4. Bowers, S.M., Sengupta, K., Dasgupta, K., Parker, B.D., Hajimiri, A.: Integrated self-healing for mm-wave power amplifiers. *IEEE Trans. Microw. Theory Tech.* **61**(3), 1301–1315 (2013)
5. Yun, S.B., Kim, Y.J., Dong, S.S., Lee, C.H.: Hardware implementation of neural network with expandible and reconfigurable architecture. In: Proceedings of the 9th International Conference on Neural Information Processing (ICONIP'02), vol. 2, pp. 970–975
6. Funahashi, K.I., Nakamura, Y.: Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Networks* **6**(6), 801–806 (1993)
7. van Keulen, E., Colak, S., Withagen, H., Hegt, H: Neural network hardware performance criteria, neural networks, 1994. In: IEEE World Congress on Computational Intelligence, IEEE International Conference, vol. 3, pp. 1955–1958 (1994)

Dynamic Scenarios in Embedded System Verification

Hara Gopal Mani Pakala, P.L.H. Varaprasad, K.V.S.V.N. Raju
and Ibrahim Khan

Abstract Scenario-based approaches have become attractive for designs through testing of embedded systems (ES). For verification, the system under development should be analyzed during its operation. Modeling of *dynamic scenarios* provides advantages of combining runtime verification with test generation and supports dynamic analysis of systems. This paper presents application of *dynamic scenarios* for test generation of a practical embedded signal processing system as case study.

Keywords System verification · Embedded signal processing system · Dynamic evaluation · Dynamic scenarios · Test generation

1 Introduction

Verification issues of embedded systems (ES) are actively being pursued by the research community [1, 2]. Scenario-based approaches have become attractive in several aspects of system design. Examples are testing [3], design using application

H.G.M. Pakala (✉)

Electronics and Communication Engineering, Matrusri Engineering College,
Saidabad, Hyderabad, Telangana 500059, India
e-mail: gopalmaniph@yahoo.com

P.L.H. Varaprasad · I. Khan

Instrument Technology, Andhra University College of Engineering,
Visakhapatnam, Andhra Pradesh 530003, India
e-mail: plhpuvvada@yahoo.co.in

I. Khan

e-mail: profibkhan@yahoo.co.in

K.V.S.V.N. Raju

Anil Neerukonda Institute of Technology and Sciences,
Sangivalasa, Visakhapatnam, Andhra Pradesh 531162, India
e-mail: kvsrn.raju@gmail.com

scenarios [4], and system design validation [5]. This paper proposes *dynamic scenarios* for test generation as they can combine advantages of *runtime verification and testing*.

Section 2 describes related works. Section 3 brings out the architectural requirement for combining runtime verification and testing. Dynamic scenarios and proposed procedure for test case generation are described in Sect. 4. The example case study of embedded signal processing system is illustrated in Sect. 5. Typical test sequence generated based on dynamic scenarios and expected outputs are shown.

2 Related Works

The verification and validation methodologies reported may be broadly divided into (a) language-dependent [6, 7] (b) specific tool (development framework) [8] and (c) commercially available tools [9]. They are not easily accessible to embedded community, and also such tools are generally expensive. A model-based scenario testing is proposed for validation testing of nuclear safety system software [3]. The model is derived from the domain-specific user's needs. It is used to generate validation tests. These tests can provide testing coverage better than random testing [3]. The authors [5] presented a scenario-based method for system-level design validation that complements an existing model-driven design method for ES based on the SystemC UML profile. A validation tool integrated into the existing model-driven codesign environment is also presented. The present work is similar to the scenario-based method [5] for system-level design validation developed for embedded systems, but uses dynamic scenarios.

3 Methodology for Verification of Embedded System

Verification is evaluation of whether or not a product, service, or system complies with a regulation, requirement, specification, or imposed condition. Though test generation techniques are being proposed for verification problems [10], testing alone as a means of verifying is rarely sufficient and cannot guarantee correctness. For verification, detailed observations of the ES execution behavior are checked at runtime against properties that specify the intended system behavior. Runtime verification complements testing through the use of monitors derived from properties, while testing typically relies on the execution of test cases.

Model-based testing makes use of models of the system under development, especially designed to generate test cases. These models typically contain information specifying the system's intended or implemented behavior, or both. Thus, model-based testing and runtime verification are very similar, differing principally in how the behaviors are produced. As such, by proper selection of model, the advantage of combining runtime verification with testing can be obtained. The ES

under development (prototype) itself must be instrumented to report events that are relevant for monitoring, and this instrumentation is based on the architecture of the monitoring method [11]. This paper focuses on modeling issues for achieving the advantage of combining testing with runtime verification using dynamic scenarios.

4 Dynamic Scenarios Methodology

Scenarios describe the interaction among system components, the environment, and users to provide intended system-level functionality. The main scenario categories during runtime situation are data flow driven and event driven. Scenarios can also be classified as intra-task scenarios, which appear within a sequential part of an application (i.e., a task), and inter-task scenarios for multitask applications. This classification can be seen as a hierarchy, and they also belong to the two former categories only. Data flow-driven scenarios characterize different executed actions in an application that are selected at runtime based on the input data characteristics (e.g., the type of object). Event-driven scenarios are selected at runtime based on events external to the application, such as user requests or system status changes (e.g., mode change). They typically characterize different quality levels for the same functionality, or performance feature, which may be implemented as different algorithms or different quality parameters in the same algorithm. These runtime scenarios are denoted as *dynamic scenarios*. They are subset of *system scenarios* [4] that describe the intended behavior of the system.

After scenario identification, they are clustered based on their classification and also on the similarities of input data characteristics or user requests or system status changes, into dynamic scenarios. It is to be noted that test cases generated based on dynamic scenarios will be a subset of the data generated based on input domain modeling of the system under development. This relation is shown in Fig. 1. This will reduce the number of test cases required to be generated for verification purposes.

4.1 Test Case Generation Using Dynamic Scenarios

The domain models are more useful for ES projects with partial requirements. The first step is to build domain-specific operations (DOP) model that characterizes the important properties of the domain from the specifications/requirements, both for development and verification. The two important requirements of the DOP model are the following:

1. Contain all the information necessary to model an application, and observation models for understanding the state of the system in uncertain environments.
2. Independent of an application and support the development of multiple applications in a given application domain.

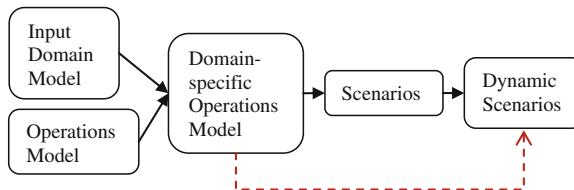


Fig. 1 Domain-specific operations model and dynamic scenarios

If the requirements themselves are not clear, then the model itself may be used in firming up the requirements/specifications iteratively. The approach for introducing dynamic scenarios into the proposed design verification path consists of three steps that are influenced by the operation environment. They are (a) identification of scenarios, (b) classification and grouping into minimal set, and (c) test case generation.

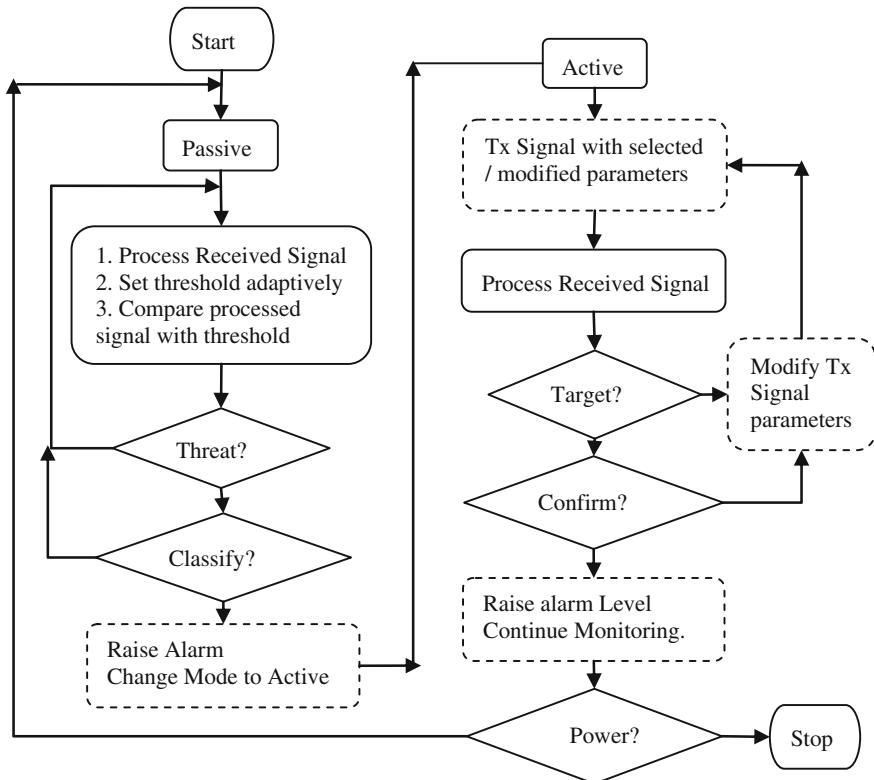
It is to be noted that when the requirements/specifications are partial, the first identified scenario itself can be used as “dynamic scenarios” for which input and expected outputs are generated. Alternately, data association among the interacting physical objects defines complete input domain for the application, and based on functionality, dynamic scenarios can be traced back. Describing input domain of the application analytically provides another way of generating test cases.

5 Case Study: Harbor Acoustic Surveillance System

A practical example Harbor Acoustic Surveillance System (HASS) [12] is considered which is a distributed sensor network comprising of several nodes. It is assumed that each node operates in a designated sector to “observe the environment, detect threat object,” and transmit “information on the threat” to a central decision-making system.

5.1 Operational Issues and Identification of Scenes

The Harbor Surveillance System’s most important subsystem is Multi-mode Detection Subsystem (MDS) [13]. Based on architecture, it will be operating in each node and “observe the environment, detect threat object, and raise the alarm”. Its operation flow chart is shown in Fig. 2. For the MDS, “passive mode” should be default operation condition or it must receive a “command” for initiating the operation. Table 1 lists typical scenes for evaluating the HASS based on MDS operation flow.

**Fig. 2** System operation flow chart**Table 1** Typical dynamic scenes

S. no.	Content	MDS function	Mode	Result
Scene 1	Ambient noise level 1	Noise analysis	Passive	Set alarm?
Scene 2	Ambient noise level 2	Noise analysis	Passive	Change mode
Scene 3	Ambient + reverberation	Correlation/spectrum	Active	Threat?/change mode
Scene 4	Ambient + reverberation + echo (range max)	Correlation/spectrum	Active	Threat evaluation
Scene 5	Ambient + reverberation + echo (<range max)	Correlation/spectrum	Active	Threat confirmation

The “Observe the Environment” scene is a complex scene with the possibility of several external events driving the scene over a period depending on the domain. Each event drives the system from one mode of operation to the other and as such requires complete input domain modeling [14, 15]. Interested readers may refer [13] for detection algorithm and [15] for more information on modeling and simulation.

5.2 Typical Test Case Generation

Table 1 lists typical dynamic scenes for evaluation. These scenarios' type and use in testing/evaluation are discussed below:

1. Scene 1: This represents “observation of environment and passive mode” operation. Ambient noise for two different time durations and variance is shown in Fig. 3. Quantum change in noise level forms the basis for mode change. Scene 2 difference is only in the value of noise levels.
2. Scene 3: The ambient noise plus reverberation of signal transmitted is shown in Fig. 4. This represents beginning of mode change and accordingly dictates

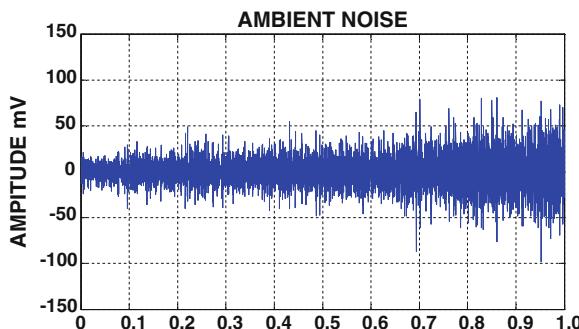


Fig. 3 Ambient noise—for 1.0 s

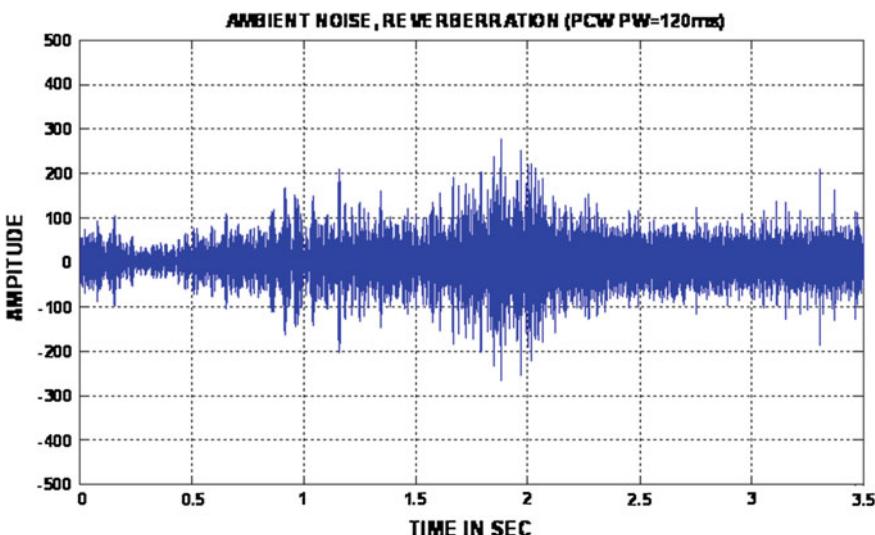


Fig. 4 Ambient noise plus reverberation PCW, ($PW = 120$ ms) no TVG

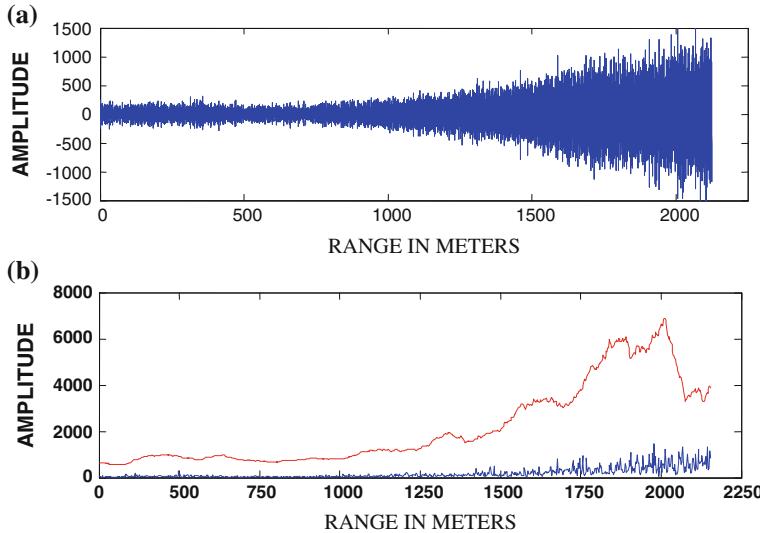


Fig. 5 Replica correlator output and threshold computed using CA-CFAR. **a** Ambient noise, reverberation ($LFM\ PW = 120$ ms) with TVG. **b** Correlator output in *blue* + threshold computed in *red*

processing requirements [13] of MDS. Normally, input data received by the HASS Sensor Array Electronics are subjected to time-varying gain (TVG) [13], and Fig. 5a shows ambient noise plus reverberation with TVG profile. If proper processing algorithm is not selected/applied, incorrect result will be given. As given in the table, correlation/spectrum analysis is employed. Correlator output (with matched replica) and threshold computed are shown in Fig. 5b. Threat detection did not take place.

3. Scene 4: “The ambient noise plus reverberation echo at given maximum range” scene, shown in Fig. 6a, is a subset of a complex scene “detect threat object”. Whether the new object that entered the “observation sector” is a “threat object” or “not threat object” is to be decided after some duration of observation. The decision involves processing, viz. passive/active signal processing [13]. The scene “processing” is data flow-driven scene and can be one of the dynamic test scenarios. Similarly, “threat object” scene can be example of external event-driven scenarios. Scenes 4 and 5 are very common during active mode of operation. The response of one popular processing algorithm called replica correlation (for a matched replica) is shown in Fig. 6b. Whenever the replica correlator output is above the computed threshold, “threat detection” is declared/confirmed.

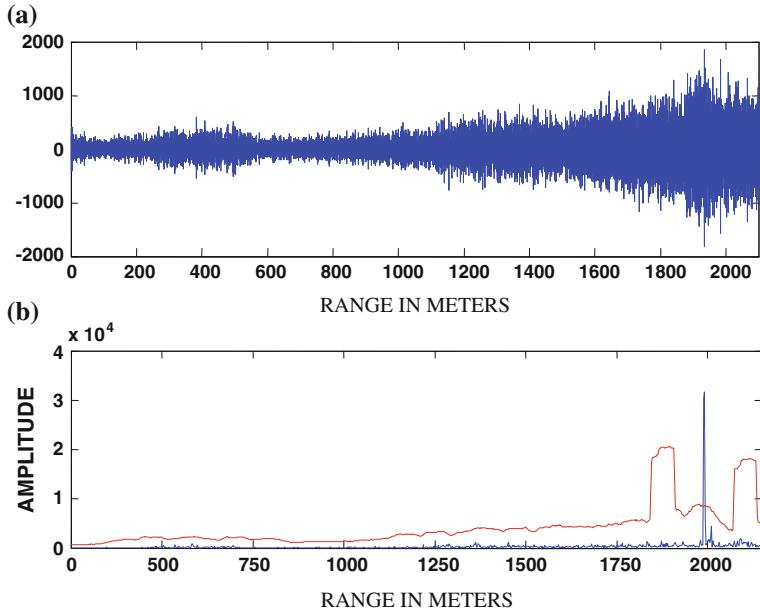


Fig. 6 a Ambient noise, reverberation (*LFM PW* = 120 ms) with TVG, ECHO at 1,900 m. b Correlator output (blue) crossing threshold computed using CA-CFAR (red) showing target detection at 1,900 m

6 Conclusions

This paper proposed test case generation based on the idea of dynamic scenarios for design verification of an embedded system application. Test cases generated based on these dynamic scenarios allow dynamic analysis of embedded systems. The advantage of dynamic scenarios is they allow generation of test cases based on input domain models. An overview of the basic steps of embedded system test case generation is described. The example case study of Harbor Surveillance System is used for illustrating the ideas presented. Test cases generated for typical scenes are described.

References

1. RTI: The economic impacts of inadequate infrastructure for software testing. Technical report, National Institute of Standards and Technology, May 2002
2. Helle, P., Schamai, W.: Towards an integrated methodology for the development and testing of complex systems. In: VALID 2013, The Fifth International Conference on Advances in System Testing and Validation Lifecycle (2013)

3. Tseng, W.-H., Fan, C.-F.: Systematic scenario test case generation for nuclear safety systems. *Inf. Softw. Technol.* **55**(2), 344–356, 2013. ISSN: 0950-5849. <http://dx.doi.org/10.1016/j.infsof.2012.08.016>
4. Gheorghita, S.V., Palkovic, M., Hamers, J., Vandecappelle, A., Mamagkakis, S., Basten, T., Eeckhout, L., Corporaal, H., Catthoor, F., Vandeputte F., De Bosschere, K.: System scenario based design of dynamic embedded systems. *ES Reports* ISSN: 1574-9517 ESR-2007–2006, Revised 23 June 2008, Eindhoven University of Technology Department of Electrical Engineering Electronic Systems, 30 Sept 2007
5. Gargantini, A., Riccobene, E., Scandurra, P., Carioni, A.: Scenario-based validation of embedded systems. In: Forum on: Specification, Verification and Design Languages, pp. 191–196 (2008)
6. Correaux, T., Beckerx, L.B., Bodeveix_z, J.-P., Farinesx, J.-M., Filali_z, M., Vernadatyz, F.: Verification based development process for embedded systems. Available at www.erts2010.org/Site/0ANDGY78/Fichier/..../ERTS2010_0029_final.pdf
7. Elshuber, M., Kandl, S., Puschner, P.: Improving system-level verification of systemC models with SPIN. In: Choppy, C., Sun, J. (eds.) 1st French Singaporean Workshop on Formal Methods and Applications (FSFMA 2013), vol. 31. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik (2013)
8. Johnson, K., Calinescu, R., Kikuchi, S.: An incremental verification framework for component-based software systems. In: CBSE'13, Vancouver, BC, Canada. 17–21 June 2013
9. Tatar M., Mauss, J.: Systematic test and validation of complex embedded systems. ERTS-2014, Toulouse, pp. 05–07 (2014)
10. Varma, P.: Design verification problems: test to the rescue? In: Proceedings of the IEEE International Test Conference (ITC'03) (2003)
11. Pakala, H.G.M., Dr P.L.H., VaraPrasad, Dr. Raju K.V.S.V.N., Dr. Khan, I.: Development of instrumentation for test and evaluation of a signal processing system. *J. Instrum. Soc. India*, **40** (1) 2012
12. Harbor Surveillance Systems: <http://www.dsit.co.il/siteFiles/1/84/5379.asp>
13. Waite, A.: Sonar for Practising Engineers, 3rd edn. Wiley, New York (2002)
14. Etter, P.C.: Underwater Acoustic Modeling and Simulation, 3rd edn. Spon Press, London, 4 Apr 2003
15. de Theije, P.A.M., Groen, H.: Multi static sonar simulations with SIMONA. In: Proceedings of 9th International Conference on Information Fusion, Florence, Italy, July 2006

PSO in Concentric Circular Arrays for Side Lobe Reduction with Symmetric Relocation Boundary Condition

Pragnan Chakravorty and Durbadal Mandal

Abstract It has been descriptively shown in the past that the boundary conditions in particle swarm optimizers are an important algorithmic part which restricts the particles within the solution space, thereby increasing their scope of participation in solution finding process. Usually, velocities of the particles are regulated to bring them back to solution space; here, a different fundamental approach of regulating the position of particles is taken; particles which go out of solution place are relocated inside by maintaining symmetry about the boundary. A three concentric ring circular antenna array is taken as the optimization target for the reduction of side lobes in the radiation pattern. Results show better performance of this boundary condition over other established ones in terms of quicker convergence and obtaining better optimization solution.

Keywords Symmetric relocation · Particle swarm · Concentric circular antenna array (CCAA) · Radiation pattern optimization

1 Introduction

Optimization techniques, particularly with evolutionary algorithms, have attracted remarkable attention of researchers across various disciplines. An important evolutionary technique is the particle swarm optimization (PSO); it has been proved faster in reaching the optimum solution than other standard techniques like genetic algorithm (GA). One of the vital aspects of a PSO algorithm is to confine the particles to the solution space or a boundary; different boundary conditions play a crucial role in PSO optimization of antenna arrays [1, 2]. Since its beginning [3],

P. Chakravorty (✉) · D. Mandal

Department of Electronics and Communication Engineering, National Institute of Technology Durgapur, Mahatma Gandhi Avenue, Durgapur 713209, India
e-mail: pciitkgp@ieee.org

D. Mandal

e-mail: durbadal.bittu@gmail.com

the PSO has been applied to variety of problems along with different boundary considerations [4–8] including those of antennas [1, 2, 9–15]. In pattern synthesis of antenna arrays, the solution spaces are generally multimodal for a complex structure; therefore, meta-heuristic techniques have shown better results than classical optimization techniques. A recent article on antenna optimization using boundary conditions which regulate the velocity of the particles [15] has been reported; meticulous analysis of the role of these boundary conditions over the performance of the PSO algorithms finds the scope and need for newer ones.

The boundary condition proposed here has a fundamental difference in its approach; unlike most of the boundary conditions, this does not control or regulate the velocity of the particles. The boundary algorithm here repositions the particles, which go out of bounds, back into the solution space. One of the earlier cases of position-regulated boundary was the hard boundary differentiated against soft boundary [1]; later, an effective position-regulated boundary called periodic boundary condition (PBC) was reported [2]. The various significant boundary conditions, including the one proposed here, are compared with each other. A significantly improved result is shown by the proposed boundary condition. We begin with a section which points out some important drawbacks of velocity regulation and thereby the need to bring up position-regulated conditions.

2 Velocity Regulation Vis-a-Vis Position Regulation

Though velocity regulation has been the dominant boundary condition applied to PSO, it has some intrinsic problems. For example, the maximum velocity (V_{\max}) has to be heuristically set between 10 and 20 % of the dynamic range of each dimension [4] or may be equal to that dynamic range [4]. Other techniques suggest an alternate expression of velocity with a constriction factor [5] or new velocity updating mechanism [7]. Despite enforcement of such conditions, the particles with any finite velocity may go beyond the solution space. The more modern approaches, such as absorbing or damping boundary [9, 10], may lead to unnecessary accrual of particles near the boundary and also adversely affect the particles intelligence of finding the global minimum (i.e., particularly, when the global minimum is not near the boundary). Moreover, forced minimization of the velocity [as shown in (1)] of a particle can let it conclude that it has reached its g_{best} and p_{best} .

$$v_n = \omega * v_n + c_1 \text{rand}() * (p_{\text{best},n} - x_n) + c_2 \text{rand}() * (g_{\text{best},n} - x_n) \quad (1)$$

when $x_n \rightarrow g_{\text{best},n}$ and $p_{\text{best},n}$, then $v_n \rightarrow v_{\text{minimum},n}$.

Few more conditions, described in [9, 10], assign a negative velocity to the particles which go out of solution space and hence bring them back; such conditions have to work outside the boundary and cannot immediately reposition the particles inside but would rather take some unknown iterative steps to do so. Unprecedented oscillations in the particles may also be seen when negative velocities are assigned.

Random velocities also fail to reposition the particles immediately into the solution space [8]. Assigning higher velocities can also take more iterative steps for the particles to converge. The better approach of boundary conditions shown in [9, 10] is that of invisible wall/boundary or floating boundary [6]; the problem is a significant number of particles may escape the solution space and those which escape cannot contribute to further exploration and hence reduce the chances of finding an optimal global minimum.

It may now be readily concluded that velocity regulation in some situations can give rise to unwanted results in the PSO process. The invisible boundary described in [9, 10] is almost an equivalent of floating/no-boundary [6]. Therefore, more suitable boundary conditions need to be sought. Position regulation has the advantage of immediately relocating the particles within the solution space, thus avoiding additional iterations and working beyond the boundaries. Control or regulation of the velocities of the particles needs not be done either; however, both positions and velocities may be regulated to give a hybrid boundary condition [2].

3 The Symmetric Relocation Boundary Condition

The simplest method of particle repositioning reported earlier is the case of hard boundary [1] where they are relocated right inside the boundary. The boundary coordinates are two in each dimension, that is, $x_{n,\max}$ for upper boundary limit and $x_{n,\min}$ for lower boundary limit. The intrinsic weakness of this method is its easy entrapment to a local minimum near the boundary. An uncontrolled accumulation of the particles takes place near the boundary limits of the solution space; Fig. 1 shows the condition after twenty iterations.

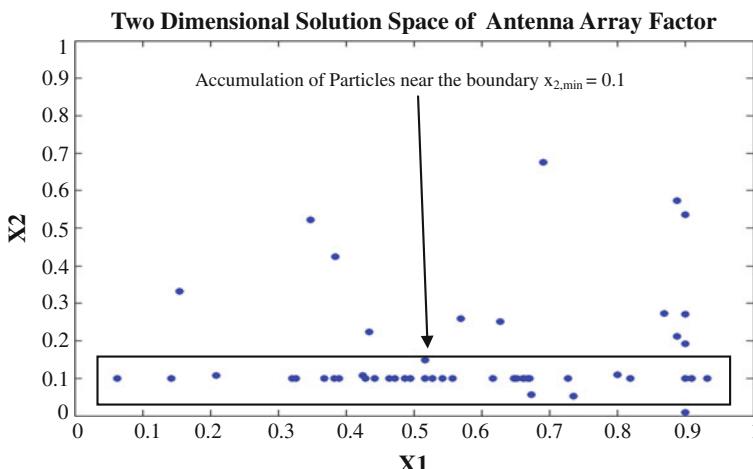


Fig. 1 Accumulation of particles near a fixed boundary in two-dimensional solution space

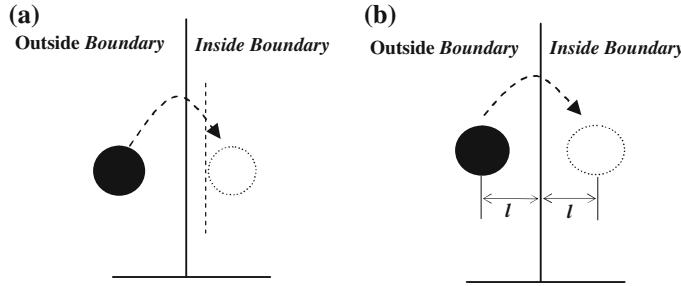


Fig. 2 **a** Hard boundary condition. **b** Symmetric relocation boundary condition

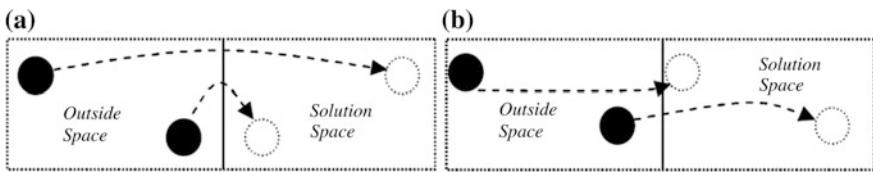
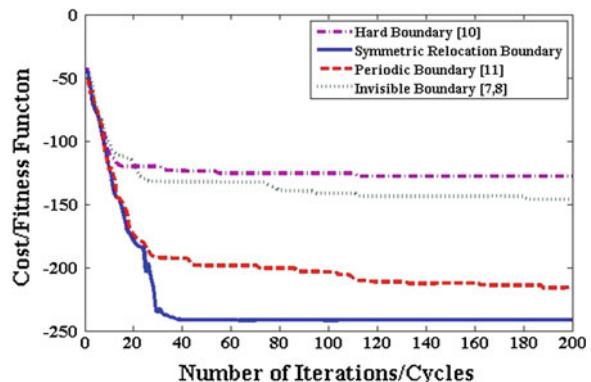


Fig. 3 Difference between **a** symmetric relocation boundary condition and **b** periodic boundary condition is shown in parts

The symmetric position relocation is a positional boundary condition technique where the particles escaping a boundary are relocated to an equidistant position ' l ' inside that boundary as shown in Fig. 2b. Here, both the lower boundary and higher boundary behave as kind of image-forming interface where the distances of the escaped particle from the respective boundaries outside the solution space are calculated; the particles are then repositioned inside the solution space with the same calculated distance about their respective boundaries. The method, though similar, is significantly different from the PBC of [2] as illustrated in Fig. 3. One of the cases of [9], namely reflected boundary condition, also sounds similar to this condition which only ensures a reversal of the velocities (i.e., a case of velocity regulation) of the particle outside the boundary. Figure 4 compares the convergence

Fig. 4 Comparison between convergence curves of different boundary conditions



curve of symmetric relocation method with PBC averaged over fifty independent runs of the particle swarm optimizer; here, the former shows better result than the later.

4 Numerical Evaluation and Results

The efficacy of proposed boundary condition is tested over a two-dimensional optimization target. The chosen dimension of optimization is attributed to its ability of enabling the visualization of moving particles over a 2D plot (i.e., Fig. 1). A three-ring concentric circular antenna array (CCAA) with variable amplitude weights and uniform spacing is taken for optimizing the reduction of side lobe levels. The elements in the inner, middle and outer rings are four (with 0.55λ inter-element spacing), six (with 0.606λ inter-element spacing), and eight (with 0.75λ inter-element spacing), respectively. The cost function is designed to work on minimization of array factor (AF) at angles or side lobes which exempt the main beam/lobe. AF for the purpose of minimization is given by [12, 13]:

$$\text{AF}(\phi, I, d_{mi}) = \sum_{m=1}^M \sum_{i=1}^{N_m} I_{mi} \exp[j(kr_m \sin \theta \cos(\phi - \phi_{mi}) + \alpha_{mi})] \quad (2)$$

where

$$kr_m = \sum_{i=1}^{N_m} d_{mi}$$

d_{mi} is inter-element spacing between the elements. I_{mi} denotes current excitation of the i th element of the m th ring, $k = 2\pi/\lambda$; λ being the signal wavelength. The angle ϕ_{mi} is the element-to-element angular separation measured from the positive x -axis. Figure 5 shows the generalized topology of CCAA. A swarm of 50 particles is taken, and the results are averaged over fifty runs.

The self-inertia ‘ ω ’ of Eq. (1) is linearly varied from 0.2 to 0.1 in these results. It is important to note that all the boundary conditions, shown in the results, do not lead to a global minimum convergence; however, if the self-inertia ‘ ω ’ is kept 0.5 or more, they all would converge to the same global minimum. The optimization is targeted at two distinct angles only (i.e., 66° and 113°) so that enough numbers of local solutions exist and the efficacy of one boundary condition over the other, to find the global solution, becomes conspicuous. Figure 4 compares the results of some significant boundary conditions. AF is plotted after optimizing the side lobe minimization target with amplitude weights found from symmetric relocation technique; the plot is compared with that of uniform amplitudes of excitation and is shown in Fig. 6.

Fig. 5 Concentric circular antenna array (CCA)

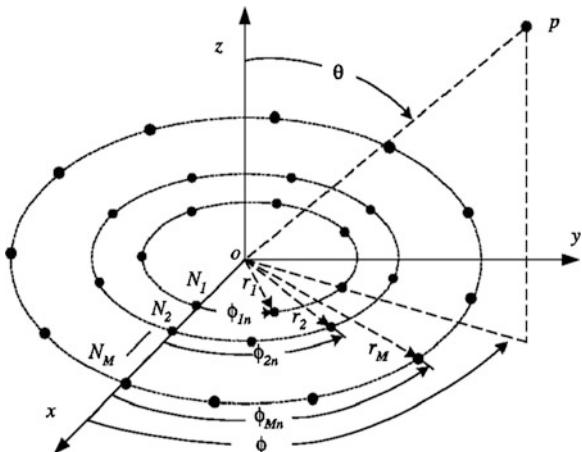
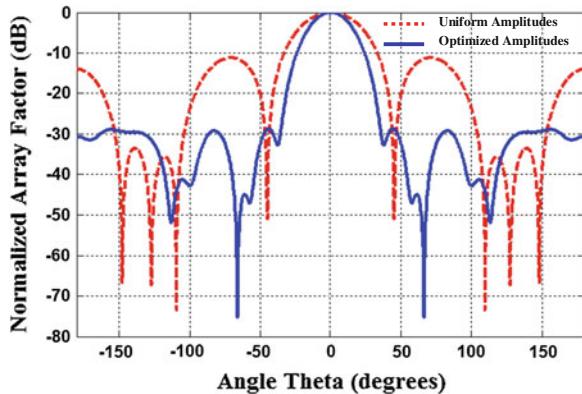


Fig. 6 Comparison of array factor plots with uniform and optimized amplitudes



5 Conclusion

The boundary conditions proposed here show significant contribution in reducing the number of iterations required in convergence of PSO process to the global minimum result. The boundary condition can also be combined with different velocity regulations to form hybrid boundary conditions. This boundary condition may also be implemented in evolutionary algorithms other than particle swarm.

Acknowledgments This work is part of a project funded by SERB, Department of Science and Technology—Government of India under the Grant SB/EMEQ-319/2013. The authors are grateful to Science and Engineering Research Board for funding the project.

References

1. Mikki, S., Kishk, A.A.: Improved particle swarm optimization technique using hard boundary conditions. *Microwave Opt. Technol. Lett.* **46**(5), 422–426 (2005)
2. Mikki, S., Kishk, A.A.: Hybrid periodic boundary condition for particle swarm optimization. *IEEE Trans. Antennas Propag.* **55**(11), 3251–3256 (2007)
3. Kennedy, J., Eberhart, R.C.: Particle swarm optimization. In: *Proceedings of the IEEE Conference on Neural Networks*, vol. 4, pp. 1942–1948. Perth, Western Australia (1995)
4. Eberhart, R.C., Shi, Y.: Particle swarm optimization: developments, applications and resources. In: *Proceedings of the 2001 Congress on Evolutionary Computation*, vol. 1, pp. 81–86 (2001)
5. Clerc, M., Kennedy, J.: The particle swarm-explosion, stability, and convergence in a multidimensional complex space. *IEEE Trans. Evol. Comput.* **6**, 58–73 (2002)
6. Galan, A.Y., et al.: Floating boundary particle swarm optimization algorithm. *Optim. Lett.* **7**, 1261–1280 (2013)
7. Bonyadi, M.R., et al.: An analysis of the velocity updating rule of the particle swarm optimization algorithm. *J. Heuristics* **20**, 417–452 (2014)
8. Li, J., et al.: A random velocity boundary condition for robust particle swarm optimization. *Lecture Notes in Computer Science*, vol. 4688, pp. 92–99 (2007)
9. Robinson, J., Rahamat-Samii, Y.: Particle swarm optimization in electromagnetics. *IEEE Trans. Antennas Propag.* **52**(2), 397–407 (2004)
10. Xu, S., Rahamat-Samii, Y.: Boundary conditions in particle swarm optimization revisited. *IEEE Trans. Antennas Propag.* **55**(3), 760–765 (2007)
11. Huang, T., Sanagavarapu, A.: A hybrid boundary condition for robust particle swarm optimization. *IEEE Antennas Wirel. Propag. Lett.* **4**, 112–117 (2005)
12. Mandal, D., et al.: Novel particle swarm optimization based synthesis of concentric circular antenna array for broadside radiation. *Lecture Notes in Computer Science*, vol. 6466, pp. 432–439 (2010)
13. Elsayid, S., et al.: Concentric circular antenna array synthesis using comprehensive learning particle swarm optimizer. *Prog. Electromagnet. Res. Lett.* **29**, 1–13 (2012)
14. Mandal, D., et al.: Wide null control of symmetric linear antenna array using novel particle swarm optimization. *Int. J. RF Microwave Comput. Aided Eng.* **21**(4), 376–382 (2011)
15. Kibria, S., et al.: A modified PSO technique using heterogeneous boundary conditions for broadband compact microstrip antenna designing. *Ann. Telecommun.* (2013). doi:[10.1007/s12243-013-0395-2](https://doi.org/10.1007/s12243-013-0395-2)

Ear Recognition Using Force Field Transform and Collaborative Representation-Based Classification with Single Training Sample Per Class

Sayan Banerjee and Amitava Chatterjee

Abstract In this paper, we propose a novel method of human identification by ear images using collaborative representation-based classification (CRC) with single training sample per class. The system first employs force field transform on each sample ear image that enhances ear structure from a redundant background, and then, the test sample is reconstructed collaboratively using training samples of all classes and eventually classified to that particular class which gives minimum reconstruction error. The presented technique achieved an encouraging recognition rate of 97.71 % with single training sample per class on a database of 330 images developed at our laboratory.

Keywords Ear recognition · Force field transformation · Collaborative representation-based classification

1 Introduction

In 1989, Iannarelli [1], a law enforcement officer in California, first proposed ear as a valid and effective biometric identifier. However, after that, for more than a decade, the researchers did not show much interest in ear as a potential biometric feature. But recently, the ear-based biometric has started gaining popularity as one of the promising modalities over the field of biometrics for its uniqueness even within twins and triplets, shape invariance capability with aging, remaining fixed over a predictable background throughout the life, and retaining its stable structure almost unchanged with facial expression variation. Similar to face as a biometric

S. Banerjee (✉) · A. Chatterjee
Jadavpur University, Kolkata, India
e-mail: sayan89.b@gmail.com

A. Chatterjee
e-mail: achatterjee@ee.jdvu.ac.in

feature, acquiring biometric data of ear from a distance makes it an appealing candidate for video surveillance too [2].

Force field transform is a comparatively new method for ear feature extraction. It was first proposed by Hurley et al. [3–5], where they used potential energy wells and channels for ear features. Later, Dong and Mu [6] successfully applied above transform for multi-pose ear feature extraction. This transform has been described in detail in Sect. 2.

Sparse representation-based classification (SRC) is a state-of-the-art, contemporary, effective classification technique that has been increasingly becoming popular in recent times and is enjoying widespread successful application in many engineering problems [7–9]. Here, each test sample is represented as a linear combination of all training samples of only that particular class to which that test sample belongs. Query image is reconstructed using all training images of different classes or subjects taken at a time via sparse coding, and it is classified to that particular class which gives minimum coding or reconstruction error [7, 8]. In the past, few researchers have explored the possible utility of SRC for the purpose of ear recognition too [10, 11].

In case of SRC, the dictionary of training images, over which the test image is coded, has to be overcomplete, which essentially means that there has to be enough training samples under each class considered [9]. But for most of the practical cases, it is not always possible or feasible to collect enough training images. Keeping those particular constraints in mind, collaborative representation-based classification (CRC) scheme has been recently proposed as a modification of SRC [8, 9] and it has been recently successfully employed for face and hand shape-based recognition problems [9, 12, 13]. Here, each test sample is represented as a linear combination of all possible training samples from all possible classes, in conformation with a basic observation that for some cases images of different classes share some similarities. This paper, to the best of our knowledge and belief, reports the first successful research attempt where force field transform has been hybridized with CRC for ear-based human recognition purposes, and the results achieved are superior than several competing methods that have been utilized for the same problem.

Our paper is organized as follows. Section 2 introduces the theory of force field transform. Section 3 presents CRC in detail. Section 4 outlines our image acquisition technique and subsequently new database formulation method using those images. Sections 5 and 6 show extensive experimental results followed by a conclusion.

2 Force Field Transform

The force field transform draws its inspiration from introductory physics, electrostatics, and electromagnetics principles. The basic theory of this transform considers that an image is transformed into force field image by assuming that each pixel is applying an attractive force to every other pixel and then replacing each pixel in the image by the corresponding force field created by rest of the pixels at that point. This

force field was referred to as Gaussian force field in [3–5]. For a 2D image I having $M \times N$ pixels, the force field at a particular pixel location (x, y) with intensity $I(x, y)$ is the vector sum of each force $f(k, l)$ applied by every other pixel of the image having location (k, l) and intensity $I(k, l)$, assuming that the original pixel is now replaced by a pixel of unit intensity. Mathematically, this force field is given by

$$FF = \sum_{k=1, \neq x}^M \sum_{l=1, \neq y}^N f(k, l) = \sum_{k=1, \neq x}^M \sum_{l=1, \neq y}^N I(k, l) \left(\frac{(k, l) - (x, y)}{\|(k, l) - (x, y)\|^3} \right) \quad (1)$$

One can see that this force field at each pixel location is a vector quantity, and hence, finally, each pixel is replaced by the absolute value of this computed quantity. In order to obtain an entire force field image, Eq. (1) is computed at each pixel location. In this context, it should be stated that if one closely follows the transform, then one can find that pixels surrounded by similar pixels will be replaced by null force fields due to cancelation of forces, which consequently highlights gradient part of the image.

3 Collaborative Representation-Based Classification

The well-known “lack of sample” problem in SRC is overcome in CRC by considering the fact that sample images of different people like ear samples share some similarities, and hence, a query ear sample y can be coded over an undercomplete dictionary A of training samples, fancying training samples from rest of the classes working as nonlocal samples to the home class of y [8]. Considering the linear system $Ax = y$ with $A \in \mathfrak{N}^{m \times n}$, $x \in \mathfrak{N}^{n \times 1}$ and $y \in \mathfrak{N}^{m \times 1}$, where A is an undercomplete dictionary with $m \gg n$, containing all training images modeled as vector, called dictionary of atoms where each of this training vector is called an atom. Relevant to our experiment, these atoms are vectored force field images. If we have c number of different classes or subjects, then A and coding vector x can be represented as $[A_1, A_2, \dots, A_c]$ and $[x_1; x_2; \dots; x_c]$, respectively, where $A_i \in \mathfrak{N}^{m \times p}$ (p denotes the number of training samples under each class) and $x_i \in \mathfrak{N}^{p \times 1}$ are submatrices and subvector associated with class i , respectively. For CRC, coding vector x is full and each of its elements suggests fraction of contribution of each training sample for reconstructing the query sample y against that of SRC, where $y \approx A_i x_i$ holds true only for y under class i , suggesting sparse nature of the coding or solution vector [8, 9, 12, 13].

CRC relies upon least square-based regularization method for solving x which proposes [8, 13]

$$x_{\text{solution}} = \arg \min_x (\|y - Ax\|_2^2 + \lambda \|x\|_2^2) \quad (2)$$

where λ is the regularization parameter. The role of λ is twofold. First of all, along with l_2 norm imposed upon the solution vector, it eradicates ill-posedness and makes the solution more stable and robust, and secondly, it induces some amount of “sparsity” to the solution [8, 13]. The above equation can be readily solved as

$$x_{\text{solution}} = Py \quad (3)$$

$P = (A^T A + \lambda I)^{-1} A^T$ plays the role of a projection matrix. As P is independent of y , it can be calculated a priori and when the query sample y comes, it projects over the sample to give the solution which subsequently reduces the classification time. Finally, the residual or reconstruction error for each class i is calculated as

$$r_i = \|y - A_i(x_{\text{solution}})_i\|_2 / \|(x_{\text{solution}})_i\|_2 \quad (4)$$

and the query sample is assigned to that class i which gives least r_i among all classes.

4 Image Acquisition and Database Creation

Right-side ear images of 22 different individuals (15 male and 7 female), with 15 samples for each subject, were captured using visual camera Nikon s2500. Each image having horizontal and vertical resolution of 300 dpi and dimension $4,000 \times 3,000$ suffers some random misalignment and motion blur of different levels due to movement of either the camera person or the subject or both. This is a practical constraint because it is almost impossible for a person to keep himself/herself stationary in a definite position while capturing his/her traits. Each RGB image was later cropped manually and transformed to grayscale format where it was



Fig. 1 Sample images from ERJU database after force field transformation

resized to 32×32 pixels to save computer memory space. The entire database has been created at Measurement and instrumentation laboratory of Electrical Engineering department of Jadavpur University, Kolkata, India, and called ERJU database for future experimentation.

Before applying CRC, each image in ERJU database was force field transformed which effectively enlightened ear structure portion, largely eliminating the effect of the redundant background, as shown in Fig. 1.

5 Experimental Result

For our experiment using ERJU database, out of 15 samples per subject, a single sample was randomly selected for training and the rest were used for testing. Each experiment was repeated over 100 times, and the mean recognition rate with standard deviation was calculated for different values of λ . The results given in Table 1 demonstrate that a best mean recognition rate of 97.71 % could be achieved with $\lambda = 0.0001$. Next, Fig. 2 shows how the performance varies when pixels were randomly corrupted by three types of noise and it can be seen that the performance never drops below 96 %. In Table 2, the performance of our system has been

Table 1 Mean recognition rate with standard deviation for different λ values

λ	Mean recognition rate (%)	Standard deviation
0.00001	97.46	0.0132
0.0001	97.71	0.0122
0.001	97.54	0.0115
0.01	97.32	0.0127
0.1	97.11	0.0129

Fig. 2 Plot of mean recognition rate of 100 trials under each experiment versus experiment number where under each trial pixels are randomly corrupted by three common types of noise at $\lambda = 0.0001$

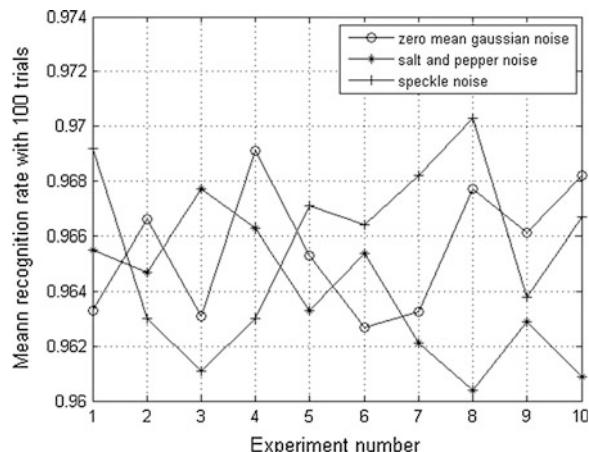


Table 2 Performance comparison with other algorithms

Method	Canny edge + SVM	PCA + SVM	Gabor feature + SRC	Our method
Mean recognition rate (%)	79.12	84.67	95.8	97.71

compared with some other state-of-the-art competing algorithms, implemented for our ERJU database. Clearly, force field transform-guided CRC outperforms other algorithms for small-sample-size problems and shows reliable performance even under the presence of noise.

6 Conclusion

A new algorithm employing force field transformation along with collaborative representation-based classification has been described for ear-based human authentication using only one training sample per subject. Experimental results show that the new algorithm demonstrates reliable performance even under presence of noise without requiring any filtration operation. One possible reason for this excellent performance can be that as force field transform performs a pixel to rest of the pixels operation and this is followed by CRC algorithm which performs regularization, the final solution arrived at may be essentially performing the noise elimination job. CRC, on the other hand, could effectively solve the small-sample-size problem too by considering extra samples from other classes as nonlocal samples. The future research will study the performance of this algorithm for an enhanced database with more number of individuals, and hybridization of CRC with other contemporary transformation algorithms will also be studied in depth.

References

- Iannarelli, A.V.: Ear Identification. Forensic Identification Series. Paramount Publishing Company, Fremont (1989)
- Yuan, L., Mu, Z.C., Yang, F.: A review of recent advances in ear recognition. In: Springer Biometric Recognition Lecture Notes in Computer Science, vol. 7098, pp. 252–259 (2011)
- Hurley, D.J., Nixon, M.S., Carter, J.N.: Force field energy functionals for image feature extraction. *Image Vis. Comput.* **20**, 311–317 (2002)
- Hurley, D.J., Nixon, M.S., Carter, J.N.: Force field feature extraction for ear biometrics. *Comput. Vis. Image Underst.* (2004)
- Hurley, D.J., Nixon, M.S., Carter, J.N.: A new force field transform for ear and face recognition. In: Proceedings of the IEEE International Conference on Image Processing, ICIP 2000, pp. 25–28 (2000)
- Dong, J., Mu, Z.: Multi-pose ear recognition based on force field transformation. In: Second International Symposium on Intelligent Information Technology Application (IITA), pp. 771–775 (2008)

7. Wright, J., Yang, A.Y., Ganesh, A., Sastry, S.S., Mai, Y.: Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(2), 210–227 (2009)
8. Zhang, D., Yang, M., Feng, X.: Sparse representation or collaborative representation: which helps face recognition? In: *IEEE International Conference on Computer Vision*, pp. 471–478 (2011)
9. Chen, X., Ramadge, P.J.: Collaborative representation, sparsity or nonlinearity: what is key to dictionary based classification? In: *IEEE International Conference on ICASSP*, pp. 5227–5231 (2014)
10. Yuan, L., Li, C., Mu, Z.: Ear recognition under partial occlusion based on sparse representation. In: *IEEE International Conference on System Science and Engineering (ICSSE)*, pp. 349–352 (2012)
11. Khorsandi, R., Cadavid, S., Abdel-Mottaleb, M.: Ear recognition via sparse representation and gabor filters. In: *IEEE Fifth International Conference on BTAS*, pp. 278–282 (2012)
12. Jia, S., Shen, L., Li, Q.: Gabor feature-based collaborative representation for hyperspectral imagery classification. *IEEE Trans. Geosci. Remote Sens.* **53**(2), 1118–1129 (2015)
13. Gangopadhyay, A., Chatterjee, O., Chatterjee, A.: Hand shape based biometric authentication system using radon transform and collaborative representation based classification. In: *IEEE Second International Conference on ICIIP*, pp. 635–639 (2013)
14. Faruqe, M.O., Al Mehedi Hasan M.: Face recognition using PCA and SVM. In: *IEEE Third International Conference on ASID*, pp. 97–101 (2009)

Demonstration of GPGPU-Accelerated Computational Fluid Dynamic Calculations

Samarth Agrawal, Manish Kumar and Somnath Roy

Abstract Problems in computational mechanics involve higher order nonlinear differential equations with complex boundary conditions, which are difficult to solve analytically and need numerical methods to predict the approximate solution. A large number of mesh points are utilized for better accuracy of the numerical technique which results in storage and operations of a large amount of data. It is of utmost importance that the time taken to perform these calculations is reduced to a realizable scale. General purpose graphical processing units (GPGPUs) provide high number of floating point operations per second (FLOPS) and potentially offer the most efficient architecture to carry out large-scale calculations in computational mechanics. In the present work, an attempt has been made to reduce the computational time for obtaining numerical solution of heat transfer by conduction, laminar flow in a rectangular channel, lid-driven cavity, and flow past square cylinder by programming GPGPUs using compute unified device architecture (CUDA), while maintaining overall second-order accuracy.

Keywords GPGPU · CUDA · Computational fluid dynamics

1 Introduction

The developments in hardware over the last few decades have enabled us to carry out large-scale scientific computations, which were previously difficult and in some cases impossible. One such development was in the field of general purpose graphical processing units (GPGPUs). Originally intended for carrying out operations pertaining to Image Processing, it can be programmed using compute unified device architecture (CUDA), a programming model, to carry out a variety of operations including matrix algebra.

S. Agrawal · M. Kumar · S. Roy (✉)

Department of Mechanical Engineering, Indian Institute of Technology Patna,
Patna 800013, Bihar, India
e-mail: somnath@iitp.ac.in

CUDA allows us to process operations in parallel on a multi-million scale. This has opened up a new debate in the field of high performance computation. While CPUs are much faster than GPGPUs for running a single set of instructions, they are capable of running programs sequentially only, i.e., they can only process one set of instructions at a time, whereas GPGPUs despite being much slower than CPUs can process millions of instructions of similar kind in parallel. Hence, serial programming is designed to decrease the time taken between two instructions (latency) and parallel programming aims at achieving higher number of jobs completed per second (throughput). Single instruction multiple data (SIMD) architecture can be efficiently developed using GPGPU to compute operations over a number of rows of a big matrix in parallel [1]. Theoretically, for a large volume of data, programs offering better throughput would achieve better computational time, however, in practice this might not be the case courtesy of several deterring factors such as copying memory from CPU to GPU and vice versa which are very expensive in terms of the time taken. Hence, it is important to an algorithm or potential parallelization considering both the volumes of numerical operations and data transfer.

Earlier efforts have been made to simulate fluid flow problems using GPU and achieved an eight-time speedup compared with a serial program over grid sizes of $3,842 \times 192$ [2]. In the present work, a different approach has been proposed that seeks to maximize throughput while keeping data transfer between the host and the device to a minimum, thus theoretically achieving a significant speedup. A study has been conducted on the effects of parallelizing programs for computing numerical solutions for three different physical phenomena: heat conduction, governed by the Laplacian equation. Channel flow, lid-driven cavity, and flow past square cylinder, governed by the Navier–Stokes equations.

2 Methodology

2.1 Heat Transfer

A steady-state two-dimensional heat conduction problem is considered. Following are the governing equations and the boundary conditions of the problem.

2.1.1 Governing Equation

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0 \quad (1)$$

2.1.2 Boundary Conditions

The left wall is considered to be adiabatic, and the top, bottom, and right walls are at 100, 0, and 50 K, respectively.

The discretized equations can be represented in the form $Ax = b$, where A is the coefficient matrix and x contains the values of temperatures at the nodes.

3 Channel Flow and Lid-driven Cavity

An unsteady, incompressible, two-dimensional, low Reynolds number flow in a rectangular channel is considered. An unsteady, incompressible, two-dimensional flow for Reynolds number = 100 in a lid-driven cavity is considered. Flow past square cylinder for low Reynolds number is solved for, which is inherently unsteady. Following are the governing equations and the boundary conditions of the problem.

3.1 Governing Equations

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = -\frac{\partial p}{\partial x} + \frac{1}{Re} \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \right) \quad (2)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} = -\frac{\partial p}{\partial y} + \frac{1}{Re} \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \right) \quad (3)$$

3.2 Boundary Conditions

Channel flow. The inlet velocities are fixed and no-slip conditions are satisfied at the walls. In addition to this, a von-Neumann boundary condition is considered at outlet.

Lid-driven cavity. No-slip condition is satisfied for all walls and the lid velocity is set at 1 m/s (upper wall).

Flow past square cylinder. No-slip condition is satisfied at the upper and bottom wall and around the cylinder. A parabolic inlet velocity profile was used, and a non-reflecting boundary condition was set for the outlet to make sure that the vortices passed the domain of interest without any perturbation [3].

The momentum and continuity equations are coupled using the Marker and cell (MAC) algorithm [4]. In addition to this, a staggered grid [4] has been employed.

4 Results and Discussion

Following are the specifications of the hardware used in the present work.

CPU—Intel(R) Xeon(R) CPU X5650 @ 2.67 GHz

GPGPU—Tesla C2075 2.0 Compute Capability.

5 Heat Transfer Problem

The resulting system of linear equations was solved iteratively using the Jacobi Method. Table 1 shows a time comparison for serial and parallel implementations. Figure 1 depicts the temperature contour. The contour lines intersect the left wall at right wall, hereby denoting the fact that no flux crosses the wall.

6 Channel Flow Problem

The time comparison of serial and parallel implementations for the channel flow case is shown in Table 2 and the resulting plot is shown in Fig. 2. The velocities near the walls tend to zero, whereas the maximum velocity is along the horizontal center line.

Case	Step size	Length	Width	Re	Mesh size	Error	Time elapsed (s)
Serial	0.02	5	1	25	250×50	10^{-4}	1,140.56
Parallel	0.02	5	1	25	250×50	10^{-4}	285.54

Table 1 Comparison of serial and parallel implementations

Case	Size	Error	Time elapsed (s)
Parallel	100×100	10^{-6}	1,985.121
Serial	100×100	10^{-6}	16,476.409

Fig. 1 Temperature contour

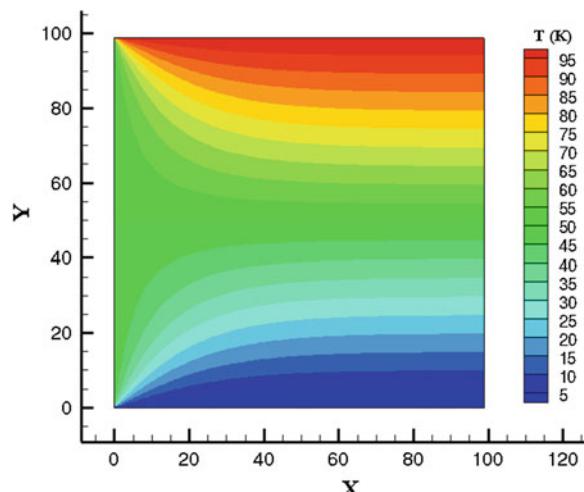
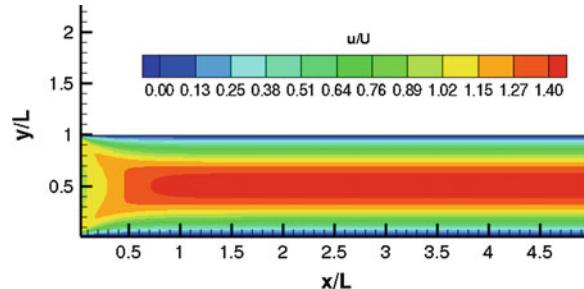


Table 2 Comparison of serial and parallel implementations for lid-driven cavity

Case	Step size	Length	Width	Re	Mesh size	Error	Time elapsed (s)
Serial	0.02	1	1	100	50×50	10^{-4}	855.05
Parallel	0.02	1	1	100	50×50	10^{-4}	114.42

Fig. 2 Non-dimensional velocity profile

7 Lid-driven Cavity

The time comparison of serial and parallel implementations for the lid-driven cavity case is shown in Table 3 and the resulting plot is shown in Fig. 3. Comparison of the horizontal velocity along the vertical center line with reported data [5] is shown in Fig. 4. In Fig. 3, one primary vortex and two secondary vortices are observed which match the reported data [6].

8 Flow Past Square Cylinder

It was observed that at very low Reynolds number ($Re = 1$), there was no separation of flow regime and Couette flow behavior exhibited, whereas for $Re = 40$, the Hoffman Bifurcation (symmetric vortices) was observed. The recirculation length

Table 3 Comparison of serial and parallel implementations for flow past square cylinder for different Reynolds numbers

Case	Step size	Length	Width	Re	Mesh size	Error	Time elapsed (s)
Parallel	0.01	6	1	1	600×100	10^{-4}	1,000.32
Serial	0.01	6	1	1	600×100	10^{-4}	6,201.984
Parallel	0.01	6	1	40	600×100	10^{-4}	1,543.67
Serial	0.01	6	1	40	600×100	10^{-4}	11,577.53
Parallel	0.01	6	1	60	600×100	10^{-4}	3,005.56
Serial	0.01	6	1	60	600×100	10^{-4}	23,443.37
Parallel	0.01	6	1	200	600×100	10^{-4}	3,300.27
Serial	0.01	6	1	200	600×100	10^{-4}	26,732.19

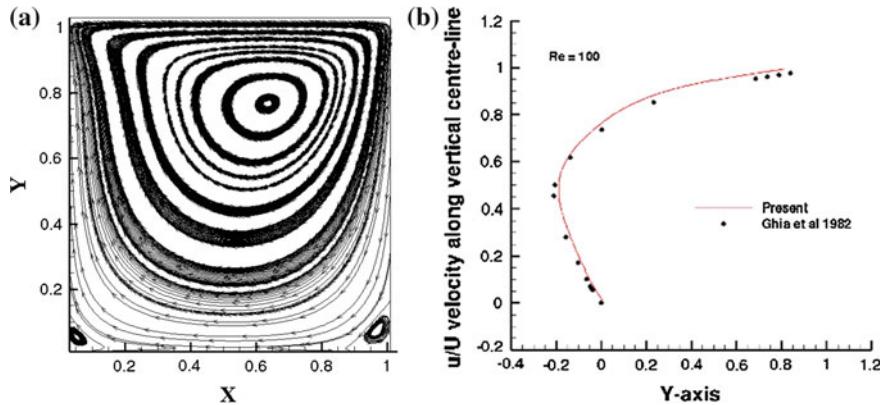


Fig. 3 **a** Streamlines, **b** comparison of horizontal velocity along the *vertical center line* with reported data⁴

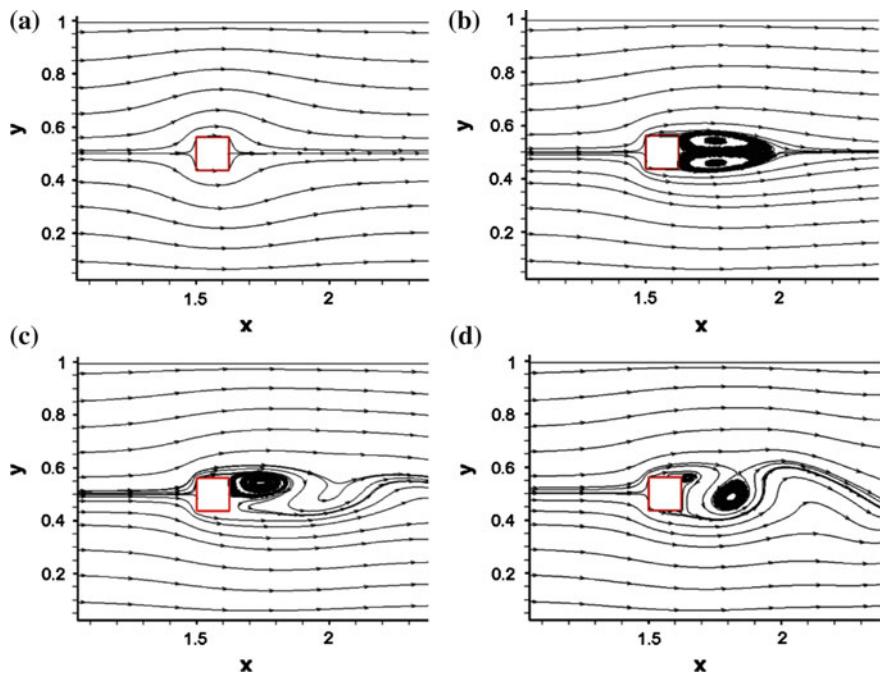


Fig. 4 Plots for flow past *square cylinder*: **a** $Re = 1$, **b** $Re = 40$, **c** $Re = 60$, **d** $Re = 200$

(x/D) was found to be 2.14 which was in agreement with Breuer et al. [7] ($x/D = 2.154$). Around $Re = 60$, we observe that Hoffman Bifurcation becomes invalid and unsymmetrical vortices (oblique) were formed, which has also been reported previously [7]. The case was also simulated for higher Reynolds number ($Re = 200$), and vortex shedding was captured.

9 Conclusion

A parallel implementation of heat transfer by conduction problem and MAC method to solve for three benchmark problems, rectangular channel flow, and lid-driven cavity has been developed by programming GPGPUs. A CUDA code was used to simulate flow past a square cylinder in a channel for different Reynolds number and an excellent agreement was found with reported data [7, 8]. A speedup has been achieved for all the cases, especially when a large number of mesh points were employed, albeit, at the expense of a higher power consumption.

Acknowledgments We thank Aeronautics Research and Development Board for sanctioning the project (Project No. R&D/SP/ME/AR&DB/IBS/2014-2015/99) and the Indian Institute of Technology Patna for allowing us to use the High Performance Cluster.

References

1. Flynn, M.: Some computer organizations and their effectiveness. *Comput. IEEE Trans.* **21**(c), 948–960 (1972)
2. Cohen, J.M., Moelmaker, M.J.: A fast double precision CFD code using CUDA. In: Biswas, R. (ed.) *Parallel Computational Fluid Dynamics: Recent Advances and Future Directions*, pp. 414–429. DEStech Publications, USA (2009)
3. Orlanski, I.: A simple boundary condition for unbounded hyperbolic flows. *J. Comput. Phys.* **21**(3), 251–269 (1976)
4. Harlow, F.H., Welch, J.E.: Numerical calculation of time dependent incompressible flow of fluid with free surface. *Phys. Fluids* **8**, 2182–2189 (1965)
5. Ghia, U., Ghia, K.N., Shin, C.T.: High-Resolutions for incompressible flow using the Navier-Stokes equations and a multigrid method. *J. Comput. Phys.* **48**(3), 387–411 (1982)
6. Muralidhar, K., Sundararajan, T.: *Computational Fluid Flow and Heat Transfer*. Narosa Publishing House, New Delhi (2004)
7. Breuer, M., Bernsdorf, J., Zeiser, T., Durst, F.: Accurate computations of the laminar flow past a square cylinder based on two different methods: lattice-Boltzmann and finite-volume. *Int. J. Heat Fluid Flow* **21**(2), 186–196 (2000)
8. Kelkar, K.M., Patankar, S.V.: Numerical prediction of vortex shedding behind a square cylinder. *Int. J. Numer. Meth. Fluids* **14**(3), 327–341 (1992)

High-performance Current Mode Receiver Design for On-chip VLSI Interconnects

Yash Agrawal, Rajeevan Chandel and Rohit Dhiman

Abstract This paper presents an efficient receiver design for on-chip current mode signaling (CMS) interconnects. The CMS interconnects using proposed receiver have 34 % lesser delay for interconnect length of 20 mm and around 3 times higher throughput at room temperature than that of conventional voltage mode signaling (VMS) interconnects. The analysis is performed for single-ended interconnects. The present work is useful for analyzing the effectiveness of voltage and current mode signaling techniques for on-chip interconnects. The simulations are performed for 180-nm technology node using Tanner EDA tool.

Keywords Current mode signaling (CMS) interconnects · Voltage mode signaling (VMS) interconnects · Delay · Power dissipation · Throughput

1 Introduction

Interconnects play a profound role in the performance of modern VLSI systems. Physical dimensions of interconnects give rise to parasitics such as resistance, inductance, and capacitance. These parasitics cause delay and power dissipation in integrated circuits. In deep submicron technologies, interconnect delay dominates over gate delay. To mitigate this issue, voltage mode buffer insertion method for interconnect systems has been reported in literature [1–3]. However, with buffer insertion, there is limit to performance improvement in terms of power and delay.

Y. Agrawal (✉) · R. Chandel · R. Dhiman

Department of Electronics and Communication Engineering, National Institute of Technology, Hamirpur 177 005, Himachal Pradesh, India
e-mail: mr.yashagrawal@gmail.com

R. Chandel

e-mail: rajeevanchandel@gmail.com

R. Dhiman

e-mail: rohitdhiman.nitham@gmail.com

Also, buffers need to be properly matched so as to have optimum results [4, 5]. Current mode signaling scheme for on-chip interconnects has been investigated and analyzed by many researchers [6–10]. In [6], Yuan has shown that CMS interconnects have an higher performance edge in terms of slew rate, bandwidth, propagation delay, supply voltage sensitivity, and electrostatic discharge (ESD)-induced damages than VMS interconnects. Also, various transmitter and receiver topologies for current mode signaling interconnects have been analyzed and presented. In [7], Dave et al. have presented variant tolerant dynamic overdriving current mode signaling scheme and have found that this scheme is robust against inter-die and intra-die variations. In [8], Kar et al. have given a closed form formula for bandwidth evaluation in CMS interconnects and have found that CMS has higher bandwidth than VMS interconnects. In [9, 10], various receiver designs for current mode signaling scheme have been presented.

In the present work, a new current mode receiver using NAND gates has been proposed. A detailed performance analysis of VMS and CMS interconnects has been carried out. The rest of the paper is organized as follows: Sect. 2 gives a brief overview of voltage and current mode interconnect systems. In Sect. 3, current mode receiver using NAND gate is presented. In Sect. 4, simulation methodology is given. The results and discussions are made in Sect. 5. Conclusions are drawn in Sect. 6.

2 Voltage and Current Mode Interconnect Systems

The voltage mode interconnect system consists of driver and receiver subsystems. The driver and receiver are implemented using CMOS inverter circuits. The voltage mode interconnect system has full voltage swing over the interconnect line. The equivalent circuit diagram of voltage mode interconnect system is shown in Fig. 1.

The current mode interconnect system consists of driver, receiver, and a decoder subsystems. The current mode driver used here is a multilevel voltage to current converter [10]. It converts 2-bit input voltage signal into 4-level current output signal. One of the advantages of this current mode driver is that 2 bits can be send simultaneously over the interconnect line. The current mode receiver senses these signals and has low input impedance termination. This low impedance termination is responsible for smaller voltage swing over the interconnect line. The small

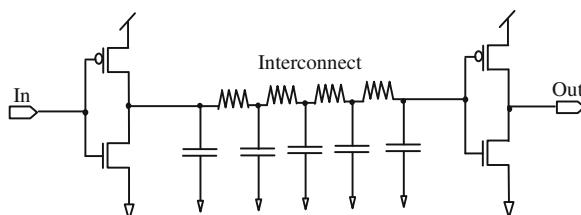


Fig. 1 Equivalent circuit of voltage mode interconnect system

voltage swing has advantage of lower latency and higher throughput in the circuit. The detailed analysis of current mode receiver subsystem is discussed in Sect. 3. The receiver output has signals in the form of thermometer codes. These thermometer codes have very less glitches since in each successive transmission of information bits, only one bit changes at a time. Also to sense and differentiate ‘10’ and ‘01’ input bit signals in multilevel current signal, thermometer codes are beneficial. The thermometer codes are then given to decoder circuit to regain the original voltage signals. The equivalent circuit diagram of current mode interconnect system is shown in Fig. 2. Its layout implementation is shown in Fig. 3.

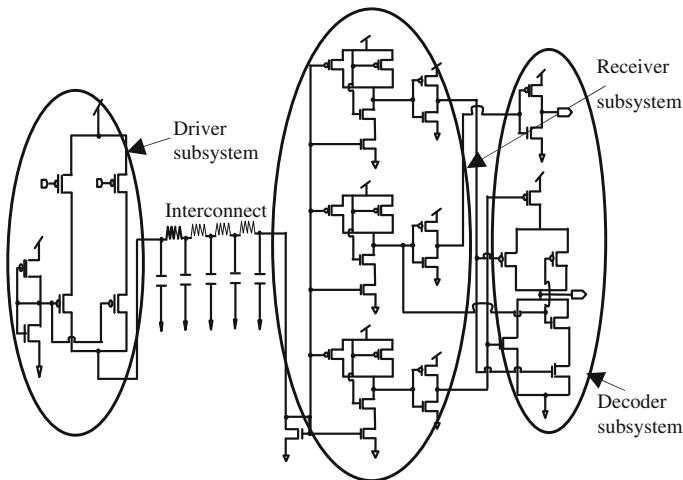


Fig. 2 Equivalent circuit of current mode interconnect system

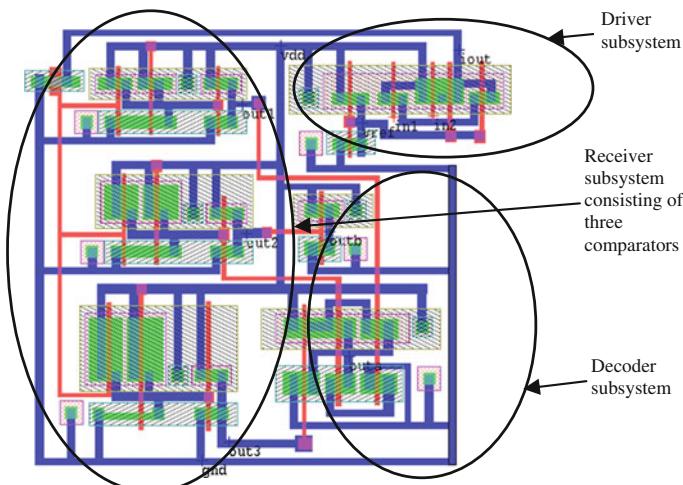


Fig. 3 Layout implementation of current mode interconnect system

3 The Receiver Circuit

The function of the receiver subsystem is to convert current signals, coming from the driver through interconnects, back to original voltage signal. This is implemented using receiver block. The receiver block consists of current to voltage converter, three comparator circuits with different threshold values, and inverters for full-swing output realization. The model of comparator block used in the receiver subsystem is shown in Fig. 4, and the receiver implementation using MOS transistors is shown in Fig. 5.

Fig. 4 Model of comparator circuit

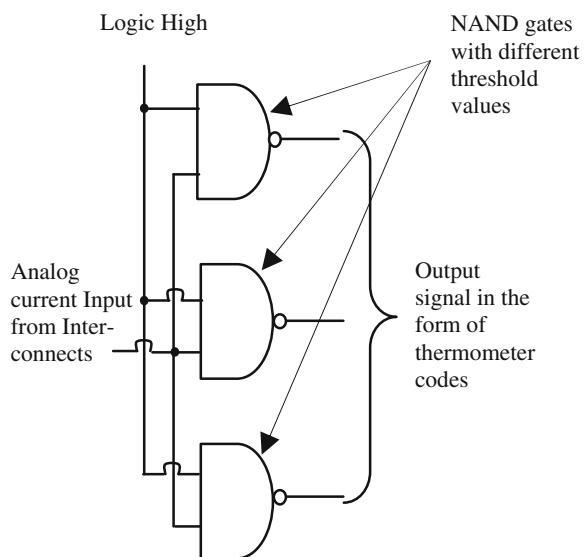


Fig. 5 MOS receiver circuit

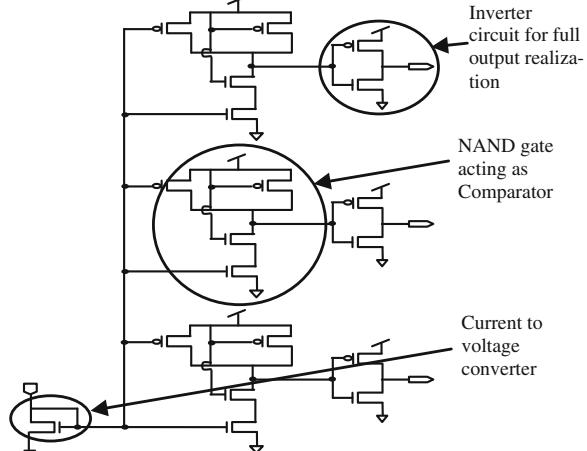
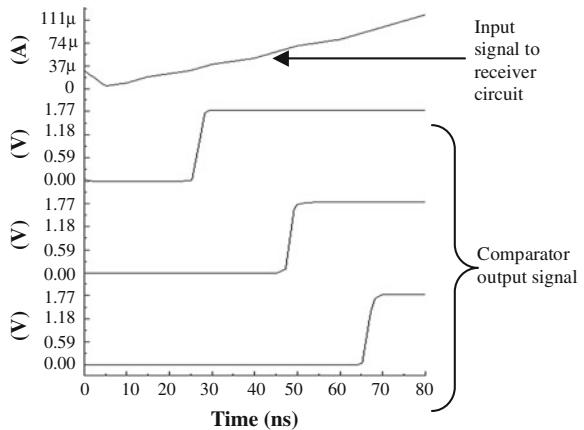


Fig. 6 Waveforms of the receiver circuit



In current mode signaling, the input impedance of the receiver circuit should be of low value. This is achieved by using a NMOS transistor whose drain and gate are shorted with each other as shown in Fig. 5. This is further input to comparator circuits. NAND gates are used for comparator design as shown in the figure. By varying width-length ratio of NMOS and PMOS transistors in NAND circuit, threshold value of the three comparators is varied. One of the inputs of NAND gate is set high, when the other input value crosses threshold value (i.e., at logic 1), then both the inputs of a NAND gate are at logic high, and the output switches to logic low. The output of different comparators switches from logic high to low value at different input current levels. This is shown in Fig. 6.

4 Simulation Methodology

The simulations are performed using Tanner EDA tool for 180 nm technology node [11]. The various constituents for computation of interconnect parasitics are as per ITRS data [12]. The values of resistance and capacitance are $0.070 \text{ M}\Omega/\text{m}$ and 270 pf/m , respectively [13]. The interconnect is modeled as four section π network. Operating temperature of 25°C and supply voltage (V_{DD}) value of 1.8 V are considered for simulations.

5 Results and Discussion

The waveforms of the current mode interconnect system are shown in Fig. 7. Here, Fig. 7a, b shows the input voltage signals. Figure 7c shows current analog signal from output of the driver subsystem. Figure 7d–f shows the thermometer codes at

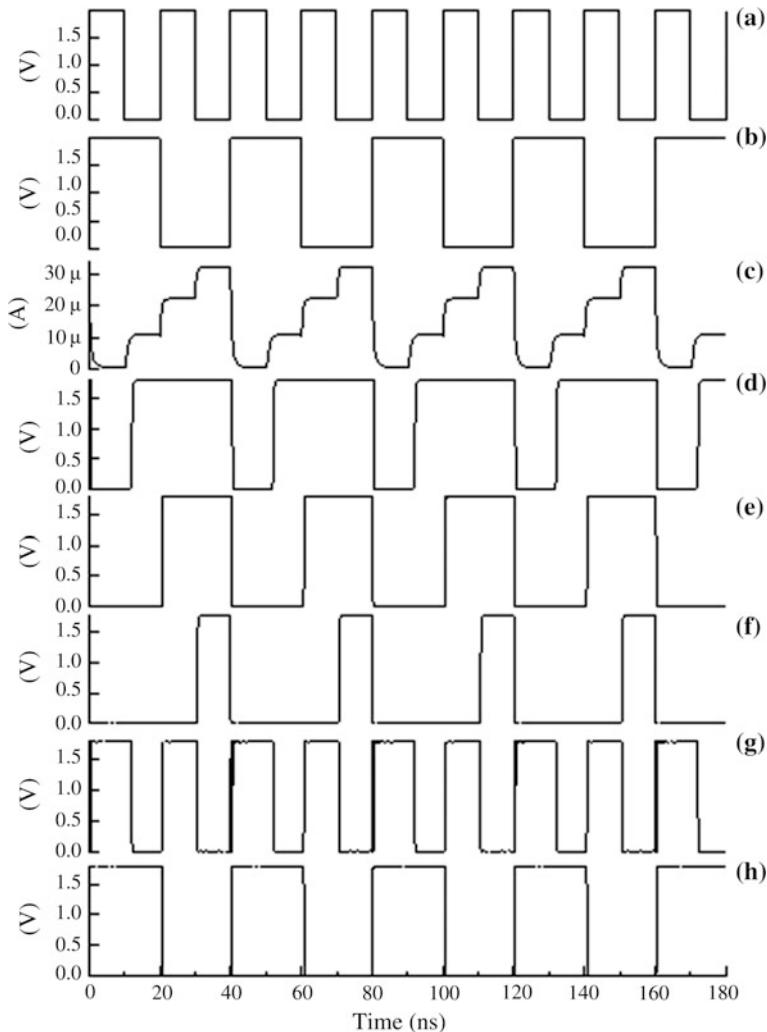
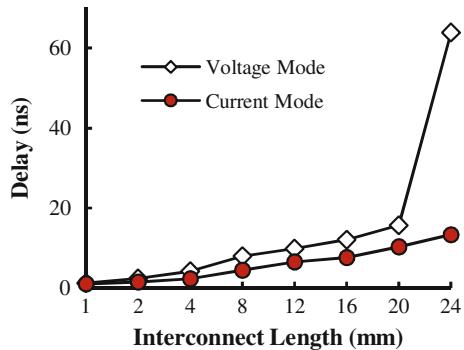


Fig. 7 Waveforms of current mode interconnect system

the output of comparator circuit. Figure 7g, h shows the original input signal regained at the output of the receiver subsystem.

The various performance parameters considered for the analysis of VMS and CMS interconnects are delay, power dissipation, and throughput. These performance parameters are investigated with the variation in interconnect length, power supply, and temperature. The interconnect length is varied from 1 to 24 mm. Power supply variations are analyzed for V_{DD} scaling from 1.8 to 1 V. The temperature variations are investigated from 0 to 80 °C. Figure 8 shows the delay variation with interconnect length. It is seen from the figure that delay in CMS interconnects is

Fig. 8 Delay variation with interconnect length



lesser than VMS interconnects. This is justified as there is less charging and discharging of parasitic capacitances in CMS interconnects as voltage swings are restricted to very small values. From the figure, it can be analyzed that at interconnect length of 20 mm, CMS has 34 % lesser latency than VMS interconnects. Figure 9 depicts power dissipation versus interconnect length plot. It is seen that for long interconnect wire lengths, power dissipation in CMS interconnects is lesser than VMS interconnects. For instance, for interconnect length of 20 mm, CMS has 7.69 % lesser power dissipation than VMS interconnects.

The variation in maximum throughput with interconnect length can be analyzed from Fig. 10. It is seen that the CMS interconnects with the proposed NAND gate receiver have higher throughput than VMS interconnects. At interconnect length of 1 mm, the maximum throughput in voltage mode and current mode signaling interconnects are 0.33 and 1 Gbps, respectively. Hence, information bits in current mode can be sent at faster rate than VMS interconnects. Figure 11 represents the delay variation with power supply scaling. It can be seen that delay in CMS interconnects is lesser than VMS interconnects for all V_{DD} values. For instance, at V_{DD} value of 1.4 V, voltage mode has 1.98 times higher delay than CMS interconnects. Power dissipation versus V_{DD} plot is shown in Fig. 12. For small values of V_{DD} , power dissipation in voltage mode is higher than CMS interconnects; however, as V_{DD} value increases, power dissipation in current mode is more than

Fig. 9 Power dissipation with interconnect length

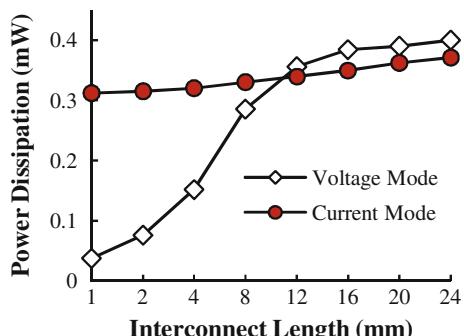


Fig. 10 Throughput variation with interconnect length

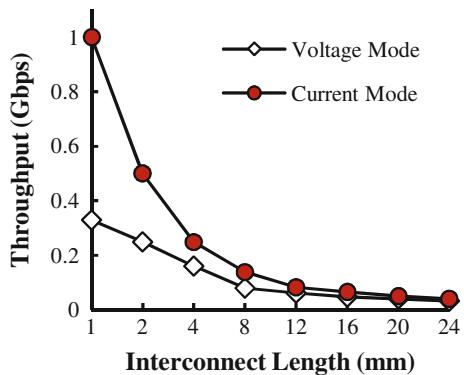


Fig. 11 Delay variation with V_{DD}

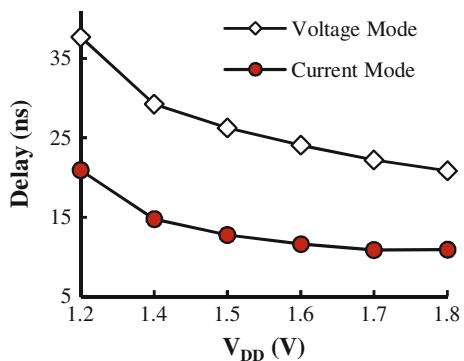
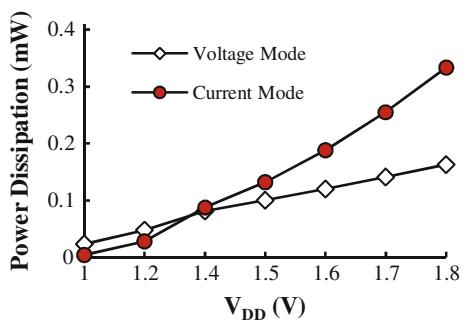
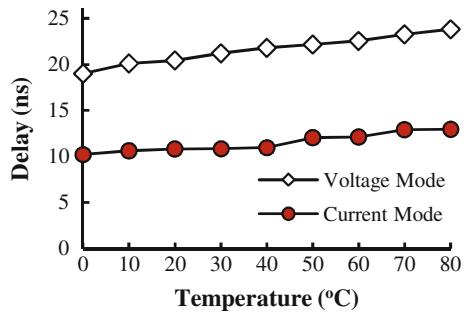


Fig. 12 Power dissipation with V_{DD}



VMS interconnects. This is because at higher V_{DD} values, large current flows through the interconnects. This causes high static power dissipation in CMS. Figure 13 shows the delay variation with temperature. From figure, it can be analyzed that there is around 45.64 % reduction in delay in current mode as compared to VMS interconnects at 80 °C.

Fig. 13 Delay variation with temperature



6 Conclusions

Design of current mode receiver using NAND circuit is proposed. The detailed analysis of both VMS and CMS interconnects is carried out. It is found that voltage mode has lesser power dissipation than CMS interconnects and is thus a suitable option for low-power applications. However, CMS interconnects have lesser delay and higher throughput than VMS interconnects. Also, voltage and temperature variations are performed, and it is seen that CMS interconnects have better performance than VMS interconnects in terms of latency. Hence, CMS interconnects would be an optimal choice for low latency and high-speed operations.

References

1. Md-Yusof, Z., Hani, M.K., Marsono, M.N., Husin, N.S.: An optimal VLSI routing algorithm with buffer insertion using iterative RLC model. In: Proceedings of IEEE International Conference on Circuits and Systems, pp. 48–53 (2012)
2. Chandel, R., Sarkar, S., Agarwal, R.P.: An analysis of interconnect delay minimization by low-voltage repeater insertion. *Microelectron. J.* **38**(4–5), 649–655 (2007). Elsevier Science
3. Dhiman, R., Chandel, R.: Sub-threshold delay and power analysis of complementary metal-oxide semiconductor buffer driven interconnect load for ultra low power applications. *J. Low Power Electron.* **8**(1), 39–46 (2012)
4. Ismail, Y.I., Friedman, E.G.: Effect of inductance on the propagation delay and repeater insertion in VLSI circuits. *IEEE Trans. Very Large Scale Integr. Syst.* **8**(2), 195–206 (2000)
5. Naeemi, A., Davis, J.A., Meindl, J.D.: Analysis and optimization for coplanar RLC lines for GSI global interconnections. *IEEE Trans. Electron Devices* **51**(6), 985–994 (2004)
6. Yuan, F.: CMOS Current Mode Circuits for Data Communication. Springer, Heidelberg (2007)
7. Dave, M., Jain, M., Baghini, M.S., Sharma, D.: A variation tolerant current-mode signaling scheme for on-chip interconnects. *IEEE Trans. Very Large Scale Integr. Syst.* **21**(2), 342–353 (2013)
8. Kar, R., Reddy, K.R., Mal, A.K., Bhattacharjee, A.K.: An explicit approach for bandwidth evaluation of on-chip VLSI RC interconnects with current mode signaling technique. In: 2nd International Conference on Computing, Communication and Networking Technologies, pp. 1–4 (2010)

9. Agrawal, Y., Chandel, R., Dhiman, R.: Design and analysis of efficient multilevel receiver for current mode interconnect system. In: 2014 IEEE Students' Conference on Electrical, Electronics and Computer Science, Bhopal, pp. 1–6 (2014)
10. Joshi, S., Sharma, D.: A novel low power multilevel current mode interconnect system. In: Emerging VLSI Technologies and Architectures (ISVLSI'06), pp. 1–6 (2006)
11. Tanner EDA tool, www.tannereda.com (2014)
12. International technology roadmap for semiconductors (ITRS) (2014)
13. Wong, S.C., Lee, G., Ma, D.: Modeling of interconnect capacitance, delay and crosstalk in VLSI. *IEEE Trans. Semicond. Manuf.* **13**(1), 108–111 (2000)

CMOS Amplifier Design Using Simplified g_m/I_D Technique

Agnish Mal, Ashis Kumar Mal and Sanjit Kumar Datta

Abstract This paper demonstrates design of analog circuits using g_m/I_D method. It explains the effectiveness of $g_m - I_D$ approach and then generates required plots using any simulator. Other than complex method of generating $g_m - I_D$ plots, which requires advanced simulators, it is shown that a plot of I_D/g_m can be easily generated directly using simulator or through a simple program capable to manipulate device current–voltage data. The success of g_m/I_D technique lies on the fact that it employs a simple rule of scaling device dimension (w) and scales the current as well as transconductance (g_m) equally, when other parameters are constant. Therefore, when a reference device with required g_m and I_D is found, it can be scaled up to generate desired g_m at the given bias current (power) I_D . Proposed method is not only technology independent, it is also free from complex mathematical expressions associated with the device as it employs data generated from simulators. As it is not based on analytical methods or models, its accuracy (independent of BSIM or ACM; and their parameter values) is much better specially for analog design. Most importantly incorporating simulator in the design process, analysis of the designed circuit, using the same simulator, is expected to match the desired performance closely. Using this simple approach, design time becomes shorter and a workable design can be made very quickly. Two basic amplifier circuits are designed using the proposed method, and simulation results are discussed.

Keywords Amplifier design · MOS sizing · g_m/I_D · Analog CMOS

A. Mal (✉)
Birla Institute of Technology Mesra, Jharkhand 835215, India
e-mail: agnish.mal@gmail.com

A.K. Mal
National Institute of Technology Durgapur, Durgapur 713209, India
e-mail: a.k.mal@ieee.org

S.K. Datta
Techno India Salt Lake, Kolkata 700091, India
e-mail: skdecenit@yahoo.co.in

1 Introduction

Analog designs are carried out traditionally by analytical methods [1, 2] and then fine-tuned using a simulator to reach the final design. It is often observed that final design variables (device dimensions) differ largely from the estimated values obtained analytically. The difference in the predicted and the final design values can be reduced if complex models like BSIM are used while carrying out the design analytically. The designer often seeks for the help of another simulator, optimizer or programming method to reach the specifications [3, 4]. Whatever may be the approach, the net time to design any analog block increases, specially for designs using short-channel devices. Moreover, such methods increase the design complexity and make the task difficult for beginners designing analog circuits in submicron CMOS.

To overcome this challenge, a technology-independent approach is proposed by Kaspers [5–7] where g_m/I_D ratio is used to as key design parameter. It is found that this ratio-based technique can be applied to all kind of devices, both long and short channel [8]. Being free from complex analytical models, even novice designers can design analog blocks with desired performance within a short span of time with fewer iteration.

In the following sections, the principle of g_m/I_D approach is explained and the method is tested by designing amplifiers using 180 nm CMOS.

2 Basic Principles

Typically, analog designers first design any analog circuit with pen and paper using some device model whose equations can be manipulated with hand calculations (SPICE level 1 or 2). Usually, these models are based on long-channel devices. Once this design is complete, the designer implements the schematic using the device sizes projected analytically on a simulator. State-of-the-art simulators, however, use extremely complex and accurate deep submicron device models like BSIM3,4 to analyse the designed circuit. Consequently, simulation results do not match with the mathematical expectations. It may be mentioned that this error is primarily due to the modelling of short-channel devices with long-channel equations. The designer now adopts an ad hoc approach, adjusts the device dimensions, bias and attempts to meet the desired response from the circuit. In CMOS technology, keeping all the transistors in saturation is one of the most challenging tasks, specially in cascaded structures.

This work investigates the use of g_m/I_D model in the design of analog circuits through simulator. Simplified circuit technique is used to generate I_D/g_m plots, so that any basic SPICE like simulator can be employed. Plot of device output impedance versus length is also generated to estimate minimum geometry of the

required transistor. Finally, simple programming-based design is discussed to make an optimized design with improved accuracy. The design process comprises of the following steps.

2.1 Theory

The current–voltage relationship of a MOS can be written as:

$$I_D = k' \cdot \frac{w}{l} f(V_{GS}, V_{DS}, \dots, V_T) \quad (1)$$

Above equation has two important characteristics. One, drain current I_D can be scaled by scaling the device width (w) if other parameters remain constant and it is independent of technology node. This simply means, maintaining bias and other parameters constant, and connecting devices in parallel will simply add their drain currents. Second, transconductance of the device ($g_m = \delta I_D / \delta V_{GS}$) is found to be:

$$g_m = k' \cdot \frac{w}{l} f'(V_{GS}, V_{DS}, \dots, V_T) \quad (2)$$

which is gain found to be scalable, provided other parameters are constant. In other words, if we connect N devices with identical bias condition but of different width w_i in parallel, overall g_m of the composite device can be written as

$$g_m = g_{m1} + g_{m2} + \dots = \sum g_{mi} \quad (3)$$

Above two points can be summarized as follows. When multiple devices of identical gate length are connected in parallel, both current and transconductance scale up (down) equally. Dividing (1) by (2) we get,

$$\frac{I_D}{g_m} = \frac{f(V_{GS}, V_{DS}, \dots, V_T)}{f'(V_{GS}, V_{DS}, \dots, V_T)} \quad (4)$$

And it is independent of device size and becomes function of bias voltages applied across different terminals. In other words,

$$\left. \frac{I_D}{g_m} \right|_{w1} = \left. \frac{I_D}{g_m} \right|_{w2} = \left. \frac{I_D}{g_m} \right|_{wN} = \phi(V) \quad (5)$$

To elaborate this issue, we have calculated g_m/I_D using α power model [9–11], where drain current is expressed as

$$I_D = k' \frac{w}{l} (V_{GS} - V_T)^\alpha \quad [\alpha = 1.3 - 2] \quad (6)$$

Therefore, $g_m/I_D = \alpha/(V_{GS} - V_T)$, clearly depends only on V_{GS} , where α , V_T could be treated as constant.

To understand its implication, consider the I_D - V_{GS} plot of a reference device shown in Fig. 1. We are required to find the relative width of a device which will operate at a bias current of $I_D(\max) = I_{D2}$ and offer a $g_m = g_{m2}$ which could provide desired amplification. Therefore, we are required to find out the operating point of the reference device where it will have an $I_D = I_{D1}$ with $g_m = g_{m1}$ such that when device is scaled by k , the scaled device ($w_2 = kw_1$) will operate at $I_D(\max) = I_{D2} = kI_{D1}$ and offer $g_{m2} = kg_{m1}$.

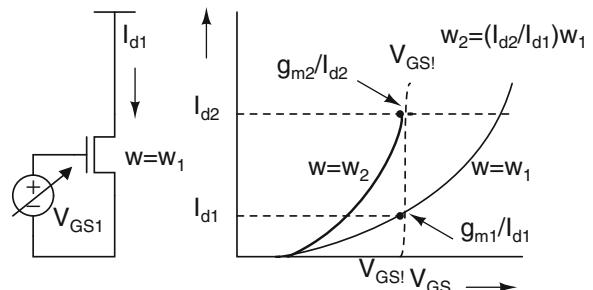
Let us summarize the steps to be followed to design an amplifier of gain A_v . From the given specifications, we need to find out (i) maximum drain current permissible (from P_D , SR, or UGF, etc.) (ii) the required load impedance R_L (from output DC bias point, output swing, etc.) then (iii) g_m required to achieve desired gain A_v and lastly calculate the ratio of drain current to required transconductance (I_D/g_m). We now take a reference device and plot its I_D/g_m ratio under varying bias. In this plot, we have to locate the bias point (I_D or V_{GS}) where desired $I_D g_m$ ratio is available. Scaling this device by k , where

$$k = \frac{I_D(\max)}{I_{D1}} \quad (7)$$

will give the dimension of the amplifying device. When I_D/g_m (or g_m/I_D) is plotted with respect to V_{GS} , we need to measure its I_{D1} at the same V_{GS} and use (7) to find the required device.

In case the ratio k is found abnormally large, which could happen if I_{D1} is small or V_{GS1} is close to V_T , designed amplifier will operate near subthreshold which may not return a reliable design. In such case, it is suggested to lower the g_m requirement and increase R_L (if permissible). Alternatively, one may use other I_D/g_m reference plots of devices of higher gate length l .

Fig. 1 Generation of (i) I_D – V_{GS} data (ii) at V_{GS1} , $g_{m1}/I_{D1} = g_{m2}/I_{D2}$



3 Generation of I_D/g_m Plots

This section describes how to generate I_D/g_m plots using simulator. Plenty of literature on this matter is available on web, and plotting technique exploits waveform manipulation tools generally available from advanced simulators like cadence. However, basic SPICE like freely downloadable simulators do not support such manipulation of plots and it becomes quite difficult to generate where I_D has to be divided by g_m under varying bias conditions.

Most preferred approach to calculate I_D/g_m ratio would be with the help of simple computer program (or SCILAB/MATLAB) which can manipulate $I_D - V_{GS}$ data easily. In this technique, simulator is required to generate a datafile where I_D as function of V_{GS} is stored by sweeping V_{GS} . Once the table of data is available, g_m can be calculated as

$$g_m = \frac{I_D(i+1) - I_D(i)}{V_{GS}(i+1) - V_{GS}(i)} \quad \text{and} \quad \frac{I_D}{g_m} = \frac{I_D(i)[V_{GS}(i+1) - V_{GS}(i)]}{I_D(i+1) - I_D(i)} \quad (8)$$

MATLAB/SCILAB can easily manipulate $I-V$ data of reference device and plot I_D/g_m as a function of either I_D or V_{GS} based on (8). Alternatively, if we wish to plot it directly in simulator, we must compute (8) while running DC sweep. To carry out the computation, we have separated 3 factors in (8). We propose to compute third factor assuming other two terms (or their ratio) are known. To appreciate this, let us rewrite (8) as follows.

$$\frac{I_D}{g_m} = \frac{(I_{D2} + I_{D1})}{2} \frac{[V_{GS2} - V_{GS1}]}{[I_{D2} - I_{D1}]} \quad (9)$$

Let us assume we can vary I_D in a manner such $I_{D1} = (1 - \delta) \cdot I$ and $I_{D2} = (1 + \delta) \cdot I$ so that $(I_{D2} + I_{D1})/(I_{D2} - I_{D1}) = 1/\delta$. Under this condition, I_D/g_m can be written as

$$\frac{I_D}{g_m} = \frac{1}{2\delta} (V_{GS2} - V_{GS1}) \quad (10)$$

If $\delta = 0.005$, I_D/g_m becomes 100 times difference between two gate-source voltages driven by current having small differential drain current. A schematic to plot I_D/g_m is shown in Fig. 2 where two controlled sources are used to drive two NMOS. Estimation of I_D/g_m could be done from differential gate-source voltages ($V_{GS1} - V_{GS2}$). One may use current ramps ($I(t) = mt \pm \delta t$) in place of controlled sources to achieve similar plot.

A better way of computing g_m/I_D is to use a *logarithmic* amplifier. In this approach, drain current (I_D) of the reference device is applied to a log amplifier and then output current is differentiated in time domain using a simple RC differentiator. Mathematically stated,

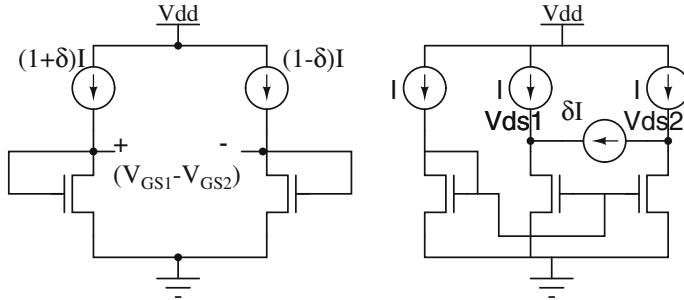


Fig. 2 Generation of (i) I_D/g_m plot (ii) g_{ds} plot for amplifier design

$$g_m = \frac{1}{I_D} \frac{\delta I_D}{\delta V_{GS}} = \frac{\delta(\ln I_D)}{\delta V_{GS}} \quad (11)$$

4 Amplifier Design

In this section, we explain the design procedure of a CS amplifier with specifications given as follows: DC gain $A_v = 10$, UGF = 1 MHz, power dissipation = 1 mW, $C_L = 5$ pF, $V_{DD} = 1.8$ V, etc. Following basic steps [2], one can find out the DC bias current and load resistance which meet the gain and other requirements. Let us assume that the required value of DC bias current is ($I_D = 20$ μ A) and $R_L = 45$ K. Ignoring g_{ds} , we get required $g_m = A_v/R_L = 222$ μ A/V. From the I_D/g_m plot (Fig. 3) of a reference device of $w/l = 240/180$ nm, we find desired I_D/g_m at $V_{GS1} = 428$ mV, at $I_{D1} = 4.42$ μ A. Thus, amplifying device needs to be scaled by 4.5x (=20/4.42) to have desired g_m at drain current of 20. Finally, when amplifier is simulated using the projected device, a voltage gain of 7.8 is measured.

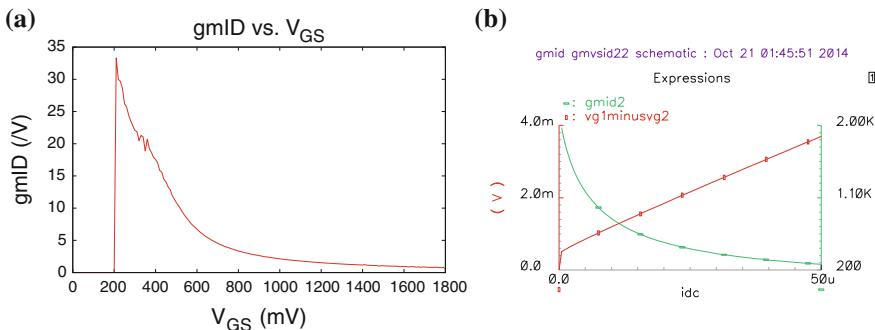


Fig. 3 **a** g_m/I_D plot using program. **b** I_D/g_m measurement using schematic

The deviation from the expected value is due to the fact that we had neglected the effects of r_o in gain calculation ($A_v \propto R_L \parallel r_o$) and approximations involved in I_D/g_m measurement. The ratio plotted using the schematic is actually $I_D/(g_m + g_{ds})$. More investigation on this is under process.

4.1 CS Amplifier with Current Source Load

Now, we turn our attention to the design of a CS amplifier with a PMOS current source load. The procedure is simple when we employ a simulator to measure the output impedance accurately. To measure g_{ds} or r_o , we use a setup where two identical devices with same gate bias are driven by $V_D = (0.5V_{DD} \pm 50\text{ mV})$ and their differential I_D (ΔI_D) is measured. Thus, $g_{ds} = 10 \times \Delta I_D$. Such plots are generated for different lengths. To find out the desired length of the load to achieve desired r_o , required length of the device is adjusted using $r_{o1}/l_1 = r_{o2}/l_2$. If r_o measured from simulation is less than the desired R_L , we need to scale up the length of the device proportionately. On the other hand, if the device r_o is higher, we expect more gain from the designed amplifier. However, one may redesign the circuit with lower current and obtain a low-power design solution. Same approach may be used if load device r_o is smaller than desired by scaling down current, provided it satisfies other specifications like slew rate.

5 Conclusion

In this paper, an extremely simple design methodology for analog circuits is discussed. Besides being technology independent, it is also free from complex mathematical expressions. Analog designers with less experience can also design opamps in a short time using the proposed approach. Being directly based on the simulator, which uses state-of-the-art MOSFET models like BSIM, the results obtained from this method are highly accurate. In this work, design of simple amplifiers for a given specification is described to validate the method. All the design formulations have been tested in UMC 180 nm CMOS process.

Acknowledgments The authors would like to acknowledge the support of VLSI Lab at NIT Durgapur, originated from SMDP project funded by DeitY, Govt. of India.

References

1. Mallya, S.M., Nevin, J.H.: Design procedures for a fully differential folded-cascode cmos operational amplifier. *IEEE J. Solid State Circ.* **24**(6), 1737–1740 (1989)
2. Allen, P.E., Holberg, D.R.: *CMOS Analog Circuit Design*. Oxford University Press, Oxford (2007)

3. Mandal, P., Vishvanathan, V.: CMOS op-amp sizing using a geometric programming formulation. *IEEE Trans. Comput. Aided Des. Integr. Circ. Syst.* **20**(1), 22–38 (2001)
4. Boyed, S., Hershenson, M., Lee, T.: Optimal design of a cmos op-amp via geometric programming. *IEEE Trans. Comput. Aided Des.* **20**(1), 1–21 (2001)
5. Silveira, D., Jespers, P.G.A.: A gm>ID based methodology for the design of CMOS analog circuits and application to the synthesis of a SOI micropower OTA. *IEEE J. Solid State Circ.* **31**(9), 1314–1319 (1996)
6. Jespers, P.: The gm>ID methodology, a sizing tool for low-voltage analog CMOS circuits. Springer, Boston (2009)
7. Rao, A.J: Analog front-end design using the gm>ID method for a pulse-based plasma impedance probe system. MS thesis, Utah State University, Utah (2010)
8. Todani, R., Mal, A.K.: Simulator based device sizing technique for operational amplifiers. *WSEAS Trans. Circ. Syst.* **13**(1), 11–28 (2014)
9. Sakurai, T., Newton, R.: Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas. *IEEE J. Solid State Circ.* **25**, 584–593 (1990)
10. Sakurai, T., Newton, R.: A simple MOSFET model for circuit analysis. *IEEE Trans. Electron Devices* **38**, 887–894 (1991)
11. Sakurai, T., Newton, R.: Alpha power-law MOS model. *Solid State Circ. Soc. Newslett.* **9**(4), 4–5 (2004)

Solution of Optimal Power Flow by an Opposition-Based Gravitational Search Algorithm

Binod Shaw, Abhik Banerjee, V. Mukherjee and S.P. Ghoshal

Abstract This paper presents an evolutionary-based approach to solve the optimal power flow (OPF) problem. The proposed approach employs opposition-based gravitational search algorithm (OGSA) for optimal settings of OPF control variables. The proposed approach is tested on the standard 26-bus test system with different objective functions those reflect minimization of fuel cost, that of transmission loss, that of sum of total voltage deviation while maintaining an acceptable system performance in terms of limits on generators' real and reactive powers, line flow limits, outputs of various compensating devices, etc. The simulation results of the proposed approach are compared with those reported in the literature. The obtained results demonstrate the potential of the proposed approach and show its effectiveness and robustness to solve the OPF problem of power system considered.

Keywords Fuel cost · Gravitational search algorithm · Opposition-based learning · Optimal power flow

B. Shaw (✉) · A. Banerjee

Department of Electrical Engineering, Asansol Engineering College,

Asansol, West Bengal, India

e-mail: binodshaw2000@gmail.com

A. Banerjee

e-mail: abhik_banerjee@rediffmail.com

V. Mukherjee

Department of Electrical Engineering, Indian School of Mines,

Dhanbad, Jharkhand, India

e-mail: vivek_agamani@yahoo.com

S.P. Ghoshal

Department of Electrical Engineering, National Institute of Technology,

Durgapur, West Bengal, India

e-mail: spghoshalnitdgp@gmail.com

1 Introduction

Optimal power flow (OPF) problem is one of the fundamental issues of power system operation and planning. Its main objective is to find the optimal settings of a given power system network that optimize certain objective function while satisfying power flow equations and inequality constraints. Several classical gradient-based optimization algorithms [1–5] have been applied to solve different OPF problems of power system. These techniques are, computationally, fast. However, conventional optimization techniques such as linear programming, nonlinear programming and Newton method have been applied to solve OPF problem assuming convex, differentiable and linear cost function [6–8]. They have the common weakness of requiring differentiable objective function, convergence to local optima and difficulty in dealing with discrete variables such as transformer tap setting and shunt capacitor bank. It is, therefore, important to develop new, more general and reliable algorithms those are capable of overcoming these drawbacks and handling such difficulties.

In view of the above, motivation of the present work arises from the application of opposition-based gravitational search algorithm (OGSA) for the solution of OPF problem of power systems. The performance of this proposed hybrid algorithm is tested on standard 26-bus test system for its OPF problem. The potential and effectiveness of the proposed approach are demonstrated by comparing the results with those reported in the recent literatures.

The rest of the paper is organized as follows. In Sect. 2, mathematical problem formulation is done. Proposed optimization algorithm and its application to the OPF problems are narrated in Sect. 3. Numerical examples and simulation results are presented in Sect. 4 to demonstrate the performance of the proposed algorithm for the OPF problems. Section 5 focuses on the conclusion of the present work.

2 Mathematical Problem Formulation

2.1 Minimization of Fuel Cost

2.1.1 Cost Function with Quadratic Cost Function

Total fuel cost of generating units having quadratic cost function without valve point effect is given by (1) and (2).

$$\text{Min } F_T = \text{Min} \sum_{i=1}^{N_G} F_i(P_i) \$/\text{h} \quad (1)$$

$$F_i(P_i) = a_i + b_i P_i + c_i P_i^2 \$/\text{h} \quad i = 1, \dots, N_G \quad (2)$$

2.1.2 Cost Function with Valve Point Effect

For more practical and accurate model of the cost function, multiple valve steam turbines are incorporated for flexible operational facilities. Total cost of generating units with valve point loading is given by (3).

$$F_i(P_i) = a_i + b_i P_i + c_i P_i^2 + |e_i \times \sin(f_i \times (P_i^{\min} - P_i))| \quad (3)$$

2.2 Minimization of Real Power Loss

The objective of the reactive power optimization is to minimize the active power loss in the transmission network, which can be defined as follows

$$\text{Min } P_{\text{Loss}} = f(\vec{X}_1, \vec{X}_2) = \sum_{k \in N_E} g_k(V_i^2 + V_j^2 - 2V_i V_j \cos \theta_{ij}) \quad (4)$$

2.3 Minimization of TVD

Minimization of TVD of load buses can allow the improvement of voltage profile. This objective function may be formulated as follows

$$\text{Min } \text{TVD} = \sum_{i \in N_L} |V_i - V_i^{\text{ref}}| \quad (5)$$

2.4 Constraints of OPF Problem

2.4.1 Equality Constraints of OPF Problem

The constraints represent the load flow equations as given below:

$$\left. \begin{aligned} P_{gi} - P_{Li} &= \sum_{j=1}^N |V_i| |V_j| (G_{ij} \cos \delta_{ij} + B_{ij} \sin \delta_{ij}) \\ Q_{gi} - Q_{Li} &= \sum_{j=1}^N |V_i| |V_j| (G_{ij} \sin \delta_{ij} - B_{ij} \cos \delta_{ij}) \end{aligned} \right\} \quad (6)$$

2.4.2 Inequality Constraints of OPF Problem

- (a) *Generator constraints*: Generator voltage, active and reactive power of the i th bus between their respective upper and lower limits and are as given by (7).

$$\left. \begin{array}{l} V_{G_i}^{\min} \leq V_i \leq V_{G_i}^{\max} \\ P_{G_i}^{\min} \leq P_i \leq P_{G_i}^{\max} \\ Q_{G_i}^{\min} \leq Q_i \leq Q_{G_i}^{\max} \end{array} \right\} \quad i = 1, 2, \dots, N_G \quad (7)$$

- (b) *Transformer tap constraints*: Transformer tap settings are bounded between upper and lower limits as represented by (8).

$$T_i^{\min} \leq T_i \leq T_i^{\max} \quad i = 1, 2, \dots, N_T \quad (8)$$

- (c) *Shunt compensator constraints*: Degree of compensation is restricted by minimum and maximum limits and is given by (9).

$$Q_{C_i}^{\min} \leq Q_{C_i} \leq Q_{C_i}^{\max} \quad i = 1, 2, \dots, N_C \quad (9)$$

- (d) *Load bus constraints*: Voltage of each PQ bus must be within its lower and upper operating limits and are given by (10).

$$V_{L_i}^{\min} \leq V_i \leq V_{L_i}^{\max} \quad i = 1, 2, \dots, N_L \quad (10)$$

- (e) *Transmission line constraints*: Line flow through each transmission line must be within its capacity limits and, mathematically, these constraints may be expressed as in (11).

$$S_{l_i} \leq S_{l_i}^{\max} \quad i = 1, 2, \dots, N_{TL} \quad (11)$$

3 Proposed Optimization Algorithm and Its Application to OPF Problem

3.1 Gravitational Search Algorithm (GSA)

Rashedi et al. propose GSA in [9]. Based on GSA, mass of each agent is calculated after computing current population's fitness given in (12) and (13)

$$m_i(t) = \frac{\text{fit}_i(t) - \text{worst}(t)}{\text{best}(t) - \text{worst}(t)} \quad (12)$$

$$M_i(t) = \frac{m_i(t)}{\sum m_i(t)} \quad (13)$$

where $\text{worst}(t)$ and $\text{best}(t)$ are defined in (14) and (15).

$$\text{best}(t) = \min_{j \in \{1, \dots, n\}} \text{fit}_j(t) \quad (14)$$

$$\text{worst}(t) = \max_{j \in \{1, \dots, n\}} \text{fit}_j(t) \quad (15)$$

Total forces applied on an agent from a set of heavier masses should be considered based on the law of gravity as stated in (16) which is followed by calculation of acceleration using the law of motion as presented in (17). Afterwards, next velocity of an agent, [as given in (18)], is calculated as a fraction of its current velocity added to its acceleration. Then, its next position may be calculated by using (19).

$$F_i^d(t) = \sum_{j \in K^{\text{best}}, j \neq i}^{N_p} \text{rand}_j \times G(t) \times \frac{M_j(t) \times M_i(t)}{R_{ij}(t) + \varepsilon} \times (x_j^d(t) - x_i^d(t)) \quad (16)$$

$$a_i^d(t) = \frac{F_i^d(t)}{M_i(t)} = \sum_{j \in K^{\text{best}}, j \neq i}^{N_p} \text{rand}_j \times G(t) \times \frac{M_j(t)}{R_{ij}(t) + \varepsilon} \times (x_j^d(t) - x_i^d(t)) \quad (17)$$

$$v_i^d(t+1) = \text{rand}_i + v_i^d(t) + a_i^d(t) \quad (18)$$

$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1) \quad (19)$$

In GSA, the gravitational constant (G) will take an initial value (G_0), and it will be reduced with time as given in (20)

$$G(t) = G_0 \times e^{-\tau \left(\frac{\text{iter}}{\text{iter}_{\text{max}}} \right)} \quad (20)$$

3.2 Opposition-Based GSA

Tizhoosh introduced the concept of opposition-based learning (OBL) in [10]. The steps of the proposed OGSA algorithm are enumerated in Fig. 1.

Step 1	Read the parameters of power system and those of OGSA and specify the lower and upper limits of each variable.
Step 2	Population-based initialization (P_0).
Step 3	Opposition-based population initialization (OP_0).
Step 4	Select N_p fittest individuals from set of $\{P_0, OP_0\}$ as initial population P_0 .
Step 5	Fitness evaluation of the agents using the objective function of the problem based on the results of Newton–Raphson power flow analysis [11].
Step 6	Update $M_i(t)$ based on (12)–(13), $best(t)$ based on (14), and $worst(t)$ based on (15), and $G(t)$ based on (20) for $i = 1, 2, \dots, N_p$.
Step 7	Calculation of the total forces in different directions by using (16).
Step 8	Calculation of acceleration by (17) and the velocity by (18).
Step 9	Updating agents' positions by (19).
Step 10	Check for the constraints of the problem.
Step 11	Opposition based generation jumping.
Step 12	Check for constraints of the problem.
Step 13	Go to Step 5 until a stopping criterion is satisfied.

Fig. 1 Pseudo code of the proposed OGSA algorithm

4 Numerical Examples and Solution Results

The proposed hybrid OGSA has been applied to solve the OPF problems of standard 26-bus test power system. The software has been written in MATLAB 2008a language and executed on a 2.63 GHz Pentium IV personal computer with 3 GB RAM.

4.1 Test System

Standard 26-bus test power system is taken for OPF study. The generators' data and cost coefficients data are taken from [12, 13]. The transmission line data and load data may be found in [13]. The generators' active powers (except slack bus-1) and generators' voltages are taken as continuous variables, while tap settings of the regulating transformers and VAR injections of the shunt capacitors are considered as discrete variables. The minimum and maximum values for the generators and load buses are taken as 0.95 and 1.05 p.u., respectively. The minimum and the maximum values for tap settings of regulating transformers are taken as 0.95 and 1.05 p.u., respectively, while those for VAR injection limits of the shunt capacitors are taken as 0 MVAR and 5 MVAR, respectively.

4.2 Minimization of Fuel Cost for Standard 26-Bus Test Power System

The proposed OGSA method is used to solve the OPF problem of test system without valve point loading. The optimal settings of control variables, fuel cost,

Table 1 Best control variable settings for fuel cost minimization objective of standard 26-bus test power system for different techniques

Control variables	EP [14]	GA [14]	MIPSO [14]	BBO [15]	GSA [16]	OGSA
P_{G-1} (MW)	388.2014	446.2860	445.5377	447.51	446.19	480.2250
P_{G-2} (MW)	126.4632	165.0258	165.0040	164.61	164.240	141.1137
P_{G-3} (MW)	300.0000	259.9621	255.8896	257.35	256.991	287.4120
P_{G-4} (MW)	150.0000	136.5176	149.5204	149.84	150.000	111.9088
P_{G-5} (MW)	200.0000	164.5045	165.4862	164.98	164.240	105.3906
P_{G-26} (MW)	111.7959	103.2680	93.8851	90.454	92.369	120.0000
V_1 (p.u.)	1.0409	1.0404	1.0273	1.0322	1.0500	0.956726
V_2 (p.u.)	1.0085	1.0174	1.0500	1.0308	1.0500	0.992195
V_3 (p.u.)	1.0211	0.9943	0.9695	1.0020	1.0038	0.950000
V_4 (p.u.)	1.0006	0.9754	1.0411	1.0035	1.0410	0.960991
V_5 (p.u.)	0.9752	1.0090	1.0063	1.0279	1.0493	0.900000
V_{26} (p.u.)	1.0249	1.0048	1.0500	1.0364	1.0500	1.058043
T_{2-3} (p.u.)	1.0800	1.0100	1.0000	1.0000	1.0500	0.915936
T_{2-13} (p.u.)	0.9900	1.0000	1.0100	0.9800	0.9900	0.903979
T_{3-13} (p.u.)	1.0000	0.9900	0.9700	0.9700	0.9600	0.979792
T_{4-8} (p.u.)	1.0200	1.0200	0.9800	1.0100	1.0000	0.900000
T_{4-12} (p.u.)	1.0600	1.0300	1.0400	0.9700	1.0100	1.050000
T_{6-19} (p.u.)	0.9900	0.9800	0.9600	0.9800	0.9800	1.049616
T_{7-9} (p.u.)	0.9300	0.9600	0.9600	0.9700	0.9700	0.900000
Q_{C-1} (MVAr)	1.0000	5.0000	4.5000	3.0000	2.5000	2.8481
Q_{C-4} (MVAr)	1.5000	3.0000	4.0000	5.0000	5.0000	0
Q_{C-5} (MVAr)	3.5000	4.0000	2.5000	4.0000	2.0000	2.9367
Q_{C-6} (MVAr)	1.5000	2.0000	3.5000	5.0000	5.0000	4.8718
Q_{C-9} (MVAr)	1.5000	4.0000	2.5000	5.0000	5.0000	0.7944

(continued)

Table 1 (continued)

Control variables	EP [14]	GA [14]	MPSO [14]	BBO [15]	GSA [16]	OGSA
$Q_{C,11}$ (MVAR)	4.0000	2.5000	3.5000	5.0000	5.0000	0.6870
$Q_{C,12}$ (MVAR)	2.0000	5.0000	2.5000	5.0000	5.0000	5.0000
$Q_{C,15}$ (MVAR)	0.5000	5.0000	2.0000	5.0000	5.0000	0
$Q_{C,19}$ (MVAR)	2.5000	2.5000	5.0000	5.0000	5.0000	4.4050
Fuel cost (\$/h)	15590	15572	15514	15488	15467.45	15462.00
P_{Loss} (MW)	13.4588	12.5594	12.3212	12.8606	16.0094	23.42
TVD (p.u.)	0.5678	0.4599	0.4346	0.3796	0.6623	0.2653
CPU time (s)	15.3001	12.7887	14.2371	13.5732	12.1098	11.00

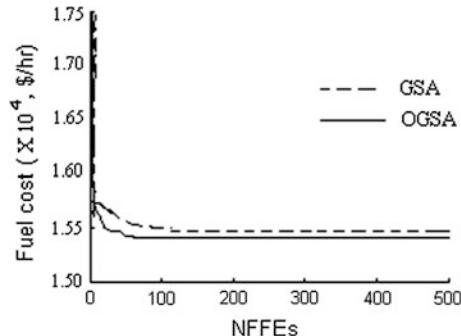


Fig. 2 Comparative convergence profiles of fuel cost for fuel cost minimization objective of standard 26-bus test power system

P_{Loss} , TVD and CPU time as reported in EP [14], GA [14], mixed integer PSO (MIPSO) [14], BBO [15] and GSA [16] and those obtained from OGSA are shown in Table 1. The simulation results presented in this table show that OGSA gives the least value of fuel cost (15462.00 \$/h) among all the other algorithms. From the results presented in Table 1, it is clear that OGSA algorithm has succeeded in finding the near global optimum solution. The comparative convergence characteristic of minimum of fuel cost as obtained by GSA and the proposed OGSA is illustrated in Fig. 2.

4.3 Minimization of System P_{Loss} for Standard 26-Bus Test Power System

The proposed OGSA is applied for P_{Loss} minimization objective of this test power network. The results reported in the literature such as GSA [16] and BBO [15] are compared to OGSA-based results, and the potential benefit of OGSA as an optimizing algorithm for this specific application is presented in Table 2. The best solutions of the minimum P_{Loss} as obtained OGSA and other aforementioned algorithms are also presented in Table 2. The obtained minimum real power loss from the proposed approach is found to be **10.3100 MW**. The value of P_{Loss} (MW) yielded by OGSA is **0.2773 MW** (i.e. 2.62 %) less than the GSA-based best result of 10.5873 MW reported in [16]. The comparative convergence profiles of P_{Loss} (MW), as yielded by GSA and the proposed OGSA, for this test power system is depicted in Fig. 3.

Table 2 Best control variable settings for P_{Loss} minimization objective of standard 26-bus test system power for different techniques

Control variables	BBO [15]	GSA [16]	OGSA
P_{G-1} (MW)	446.9439	451.7632	469.9638
P_{G-2} (MW)	200.0000	199.8991	198.8739
P_{G-3} (MW)	209.0601	206.3199	192.6165
P_{G-4} (MW)	150.0000	150.0000	149.6981
P_{G-5} (MW)	182.0582	180.7755	182.6257
P_{G-26} (MW)	86.4766	85.7769	80.2620
V_1 (p.u.)	1.0500	1.0500	1.0601
V_2 (p.u.)	1.0493	1.0494	1.0578
V_3 (p.u.)	1.0402	1.0469	1.0384
V_4 (p.u.)	1.0195	1.0214	1.0422
V_5 (p.u.)	1.0850	1.0499	1.0555
V_{26} (p.u.)	1.0500	1.0497	1.0590
T_{2-3} (p.u.)	0.9600	0.9600	0.9555
T_{2-13} (p.u.)	1.0900	1.1000	0.9546
T_{3-13} (p.u.)	0.9900	0.9900	0.9978
T_{4-8} (p.u.)	0.9900	0.9900	0.9819
T_{4-12} (p.u.)	0.9800	0.9800	0.9858
T_{6-19} (p.u.)	0.9600	0.9700	0.9694
T_{7-9} (p.u.)	0.9400	0.9500	0.9685
Q_{C-1} (MVAR)	1.0000	4.0000	1.3700
Q_{C-4} (MVAR)	4.0000	5.0000	0.6900
Q_{C-5} (MVAR)	4.5000	3.5000	2.5200
Q_{C-6} (MVAR)	5.0000	5.0000	1.7800
Q_{C-9} (MVAR)	5.0000	5.0000	3.1900
Q_{C-11} (MVAR)	5.0000	5.0000	3.2800
Q_{C-12} (MVAR)	5.0000	5.0000	3.7800
Q_{C-15} (MVAR)	5.0000	5.0000	1.5300
Q_{C-19} (MVAR)	5.0000	5.0000	0.9100
Fuel cost (\$/h)	15605.0070	15644.1368	15792
P_{Loss} (MW)	10.6615	10.5873	10.3100
TVD (p.u.)	0.7171	0.6910	0.5428
CPU Time (s)	13.7834	13.5432	12.10

4.4 Minimization of System TVD for Standard 26-Bus Test Power System

Minimization of TVD objective is considered for this case. The OGSA-based simulation results of control variables, fuel cost, P_{Loss} , TVD and CPU time for this minimization objective are shown in Table 3. The same results, as obtained by BBO

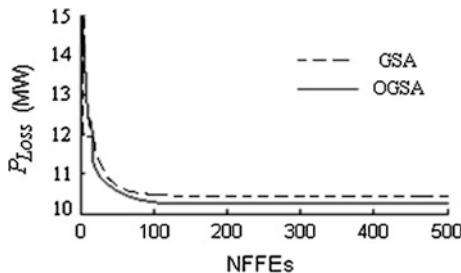


Fig. 3 Comparative convergence profiles of P_{Loss} for transmission loss minimization objective of standard 26-bus test power system

Table 3 Best control variable settings for TVD minimization objective of standard 26-bus test power system for different techniques

Control variables	BBO [15]	GSA [16]	OGSA
P_{G-1} (MW)	497.7046	494.0021	500.000
P_{G-2} (MW)	198.8158	199.6346	161.9090
P_{G-3} (MW)	181.0286	171.7571	80.0033
P_{G-4} (MW)	149.9631	147.7545	109.2449
P_{G-5} (MW)	196.7456	194.2034	199.9977
P_{G-26} (MW)	52.4986	69.6752	92.5403
V_1 (p.u.)	1.0237	1.0228	1.0317
V_2 (p.u.)	1.0215	1.0225	0.9411
V_3 (p.u.)	1.0000	1.0000	0.9942
V_4 (p.u.)	1.0000	1.0000	1.0304
V_5 (p.u.)	1.0500	1.0499	1.0481
V_{26} (p.u.)	1.0036	1.0069	1.0102
T_{2-3} (p.u.)	1.0100	0.9900	1.0135
T_{2-13} (p.u.)	1.0100	1.0800	0.9923
T_{3-13} (p.u.)	0.9800	0.9800	0.9906
T_{4-8} (p.u.)	1.0200	1.0200	1.0069
T_{4-12} (p.u.)	0.9600	0.9600	0.9660
T_{6-19} (p.u.)	0.9000	0.9000	0.9187
T_{7-9} (p.u.)	0.9800	0.9800	0.9854
Q_{C-1} (MVAR)	0.0000	5.0000	2.9700
Q_{C-4} (MVAR)	4.5000	4.5000	2.6700
Q_{C-5} (MVAR)	5.0000	3.5000	1.3900
Q_{C-6} (MVAR)	5.0000	5.0000	3.3900
Q_{C-9} (MVAR)	5.0000	4.5000	2.6600
Q_{C-11} (MVAR)	5.0000	5.0000	2.0600
Q_{C-12} (MVAR)	5.0000	5.0000	1.9700
Q_{C-15} (MVAR)	5.0000	5.0000	5.0000
Q_{C-19} (MVAR)	5.0000	5.0000	2.9400
Fuel cost (\$/h)	15732.5163	15772.3144	14146
P_{Loss} (MW)	13.5072	13.2551	14.3100
TVD (p.u.)	0.2549	0.2534	0.1418
CPU time (s)	12.5978	12.1071	11.04

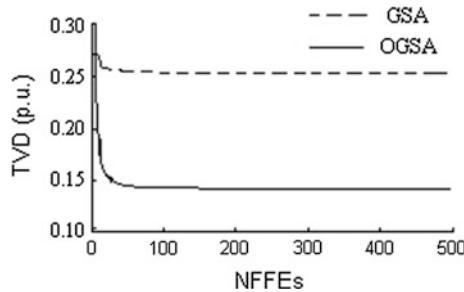


Fig. 4 Comparative convergence profiles of TVD for TVD minimization objective of standard 26-bus test power system

[15] and GSA [16], are also featured in this table. The results presented in the table show that TVD found by the proposed OGSA method (0.1418 p.u.) is lower than that offered by BBO [15] (0.2549 p.u.) or GSA [16] (0.2534 p.u.) though the value of P_{Loss} yield by OGSA is slightly higher than that obtained by GSA. Thus, OGSA gives the best voltage profile among other algorithms. The comparative convergence characteristic of this objective function, as yielded by GSA and the proposed OGSA, is plotted in Fig. 4.

5 Conclusion

In this paper, the proposed OGSA algorithm has been successfully implemented to solve OPF problem of standard 26-bus test power network. It has been observed that the proposed OGSA algorithm has the ability to converge to a better quality near-optimal solution. This algorithm possesses better convergence characteristic and robustness than other prevailing techniques reported in the recent literature. It is also clear that the proposed OGSA algorithm is free from the shortcoming of premature convergence exhibited by the other optimization algorithms. Thus, the proposed algorithm may become a promising tool for solving some other more complex engineering optimization problems for the future researchers.

References

1. Happ, H.H.: Optimal power dispatch a comprehensive survey. *IEEE Trans. Power App. Sys.* **96**(3), 841–851 (1977)
2. Bouktir, T., Belkacemi, M., Zehar, K.: Optimal power flow using modified gradient method. In: Proceeding ICEL' 2000, vol. 2, pp. 436–442. U.S.T. Oran, Algeria (2000)
3. Capitanescu, F., Glavic, M., Ernst, D., Wehenkel, L.: Interior-point based algorithms for the solution of optimal power flow problems. *Elect. Power Syst. Res.* **77**(5–6), 508–517 (2007)

4. Carvalho, E.P.D., Junior, A.D.S., Ma, T.F.: Reduced gradient method combined with augmented lagrangian and barrier for the optimal power flow problem. *Appl. Math. Comput.* **200**(2), 529–536 (2008)
5. Ramos, J.L.M., Exposito, A.G., Quintana, V.H.: Transmission power loss reduction by interior-point methods: implementation issues and practical experience. *Proc. IEE Gener. Transm. Distrib.* **152**(1), 90–98 (2005)
6. Lobato, E., Rouco, L., Navarrete, M.I., Casanova, R., Lopez, G.: An LP-based optimal power flow for transmission losses and generator reactive margins minimization. In: *Proceedings of IEEE Porto Power Tech Conference Portugal* (2001)
7. Momoh, J.A., El-Hawary, M.E., Adapa, R.: A review of selected optimal power flow literature to 1993 part I and II. *IEEE Trans. Power Syst.* **14**(1), 96–111 (1999)
8. Pudjianto, D., Ahmed, S., Strbac, G.: Allocation of VAR support using LP and NLP based optimal power flows. *IEE Gener. Transm. Distrib.* **149**(4), 377–383 (2002)
9. Rashedi, E., Nezamabadi-pour, H., Saryazdi, S.: GSA: a gravitational search algorithm. *Inf. Sci.* **179**(13), 2232–2248 (2009)
10. Tizhoosh, H.R.: Opposition-based learning: a new scheme for machine intelligence. In: *Proceedings of International Conference Computational Intel, Modeling, Control and Automation*, pp. 695–701 (2005)
11. Wang, X.F., Song, Y., Irving, M.: *Modern power systems analysis*. Springer, New York (2008)
12. Gaining, Z., Huang, H.S.: Real-coded mixed integer genetic algorithm for constrained optimal power flow. In: *Proceedings of IEEE TENCON Region 10 Conference 3*, pp. 323–326 (2004)
13. Gaining, Z., Huang, H.S.: Security-constrained optimal power flow by mixed integer genetic algorithm with arithmetic operators. In: *Proceedings of IEEE Power Engineering Society General Meeting*, pp. 1–8 Montreal (2006)
14. Gaining, Z.L.: Particle swarm optimization to solving the economic dispatch considering the generator constraints. *IEEE Trans. Power Syst.* **18**(3), 1187–1195 (2003)
15. Roy, P.K., Ghoshal, S.P., Thakur, S.S.: Multi-objective optimal power flow using biogeography-based optimization. *Elect. Power Comp. Syst.* **38**(12), 1406–1424 (2010)
16. Bhattacharya, A., Roy, P.K.: Solution of multi-objective optimal power flow using gravitational search algorithm. *Proc. IET Gener. Transm. Distrib.* **6**(8), 751–763 (2012)
17. Bhattacharya, A., Chattopadhyay, P.K.: Solution of optimal reactive power flow using biogeography-based optimization. *Int. J. Electr. Electron Eng.* **4**(8), 568–576 (2010)

Information Systems and Artificial Intelligence Technology Applied in Concrete Road Design

Erhan Burak Pancar and Muhammet Vefa Akpinar

Abstract In this paper, the main objective is to show how to obtain precise concrete road design using an information systems and artificial intelligence (ISAI) applied in the numerical design stage of a plate on elastic solid system. In this system, finite difference solution for very large plate on elastic solid is used. In this solution, computations take only few minutes using computers with very low capacity. In this paper, the logic of solution is determined according to the principles of finite difference method using an electronic interior control subsystem. In this deterministic approach, plate is divided into small parts. One of these parts, which is named intelligent object (IO), is solved under traffic loadings by ISAI technology. Then, the plate surface is scanned by IO, and precise solution for plate thickness is determined by the system. The system incorporates ISAI technology.

Keywords Concrete road · Finite difference · Intelligent object · Information systems and artificial technology

1 Introduction

The traditional methods used on concrete road design are usually based on empirical design procedures. Design catalogs have been developed in several countries, including Belgium, Germany, France, and America. There are also some computer programs to design the concrete road using finite element analysis. However, design procedure requires hundreds of thousands of stress and deflection calculations, and these computations take days to complete using existing finite

E.B. Pancar (✉)

Department of Architecture, Ondokuz Mayıs University, Samsun, Turkey
e-mail: erhanpancar@hotmail.com

M.V. Akpinar

Department of Civil Engineering, Karadeniz Technical University, Trabzon, Turkey

element programs. To reduce computer time to a practical level, neural networks, which is a statistical approach, have been developed [1–5].

A controlling objective of road design is to get a precise solution quickly, delivering an economical and safe result. This paper describes a different approach: Electronic information processing by information systems and artificial intelligence technology (ISAI) application, by which a precise concrete road design is determined. It is a deterministic approach. In this approach, a numerical solution is achieved by scanning the road surface by an intelligent object (IO) and applying ISAI technology to derive the optimum concrete road design. In this paper, logic of ISAI for concrete road design will be determined.

2 Deterministic Approach in Artificial Intelligence

In designing the ISAI to be used in this application, both deterministic and statistical methods can be suitably used to achieve the numerical solutions. This paper explains the use of a deterministic approach using ISAI in the numerical design stage of a concrete road.

An ISAI approach differs from classical numerical concrete road design in the following ways:

- The beginning plate thickness is set at the minimum of standards used in design. The system determines optimum design dimensions from this starting point.
- Deflections, bending moments, and stresses are obtained for every single point of very long road under the given traffic loadings when IO, which includes parametric properties of the road system, scans the road surface.
- The system checks all these calculations from the starting point. The plate thickness is increased by the system until the effects are balanced. The system determines the critical locations where deflections are bigger than other parts of the plate and it puts dilatations there. Then, it starts to compute according to these boundary (dilatation) conditions and optimum road thickness; dilatation locations and reinforcement are determined by the system using electronic interior control subsystem.

3 Program Design

To design the program, which is based on an ISAI application, modular and systematic design techniques are essential. The knowledge-base structure consists of program files and data files. Certain information clusters are present in the system. Each program module comes to the kernel memory as needed then flies away. As a result, all the programs and their modules that constitute the system are not in the kernel memory at the same time. The program operating system controls the

relations of the programs in the system and the relation between the data on the knowledge base and the programs.

This approach creates a platform independent of the other operating systems (Windows, Mac OS, and UNIX). This approach is also independent of the coding systems and languages used in programs. For example, some programs and their modules can be designed in different coding systems and languages to increase efficiency. When a program or its module is coded by C++, another module may be coded by FORTRAN or Pascal. The program management system described in this paper can run and manage all modules as if they are coded in one language. The coding explained in this paper is in C++ because of its efficiency and structural properties. Program can run via Internet. In Fig. 1, there is a road information system general information flow diagram. The main headings in this diagram are as follows:

- About the program (it explains ISAI technology)
- Knowledge base (standards, technical specifications, experiments, and unit prices)
- Input (information about project, geometry (not the plate thickness), material, soil, and effect)
- Numerical process (coefficient matrix, inverse matrix, and numerical solution)
- Numerical drawings (system geometry, stresses, deflections, and reinforcement details)
- Measurements (construction measurements and reinforcement measurements)
- Output (numerical process output, numerical drawing output, measurement output, and experiment output)

The system can determine the optimum concrete road design and choose economical concrete road type among jointed plain (JPCP), jointed reinforced (JRCP), and continuously reinforced (CRCP) rigid roads using ISAI technology. The system can make the solution so quick when any item of knowledge base or input changes and it is easy to enlarge the system by adding new modules for new problems.

4 Numerical Solution

The system solution is obtained by dividing general road system into the intelligent objects and solving these objects. Graphic presentation is in Fig. 2.

It is better to explain the IO to clarify the method used in this system.

Intelligent object (IO): It is the subsystem module which contains all the common properties of the system depending on the boundary conditions and values. All of the systems are solved by being scanned when this module replaces on the system depending on the system design principles. In other words, the system solution is reduced to the solution of intelligent object subsystem by using information systems and artificial intelligence technology. The difference is that a very

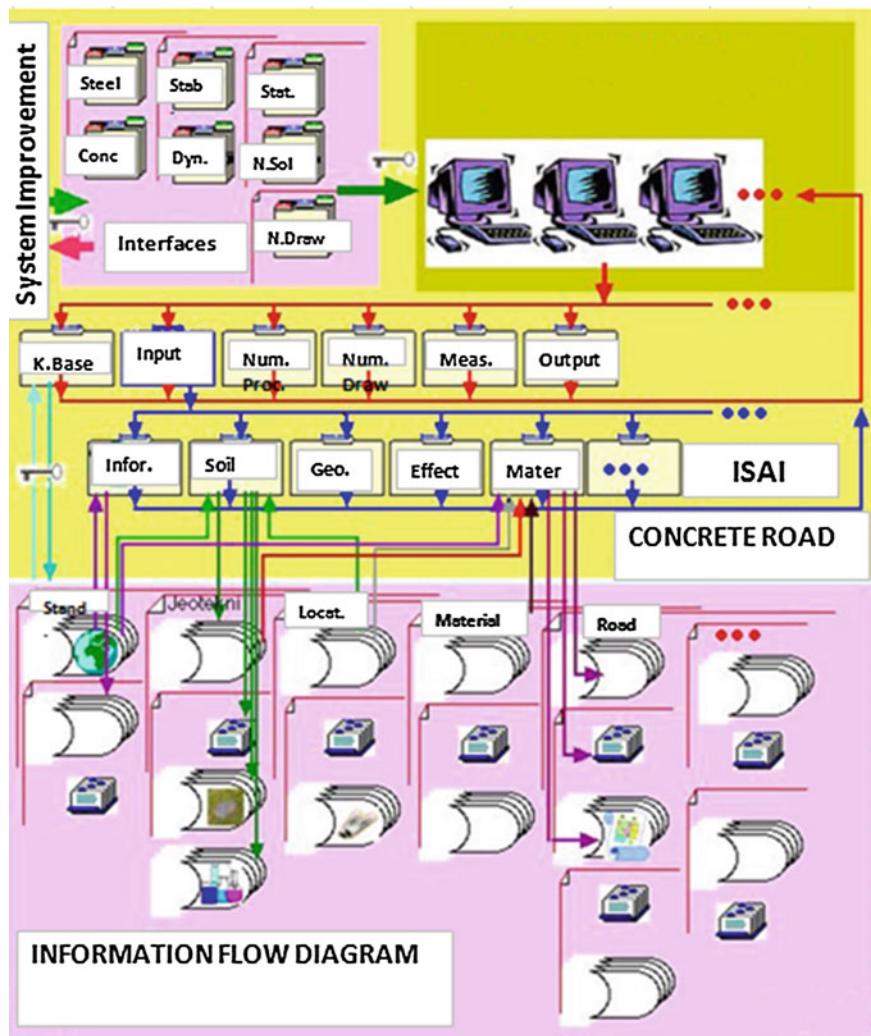


Fig. 1 Road information system general information flow diagram

large road route system is solved even by computers which have very low capacities. Therefore, it is appreciated that information systems design, which plans information flows very good, must be done.

In numerical design stage of the system, concrete road is taken as a plate on elastic solid. The program is written that uses the method of finite differences to construct the full coefficient matrix for the plate on elastic foundation. One of the most important points in this numerical solution is plate does not need to be rectangular. All nodes on the plate can be determined on infinite plate and the

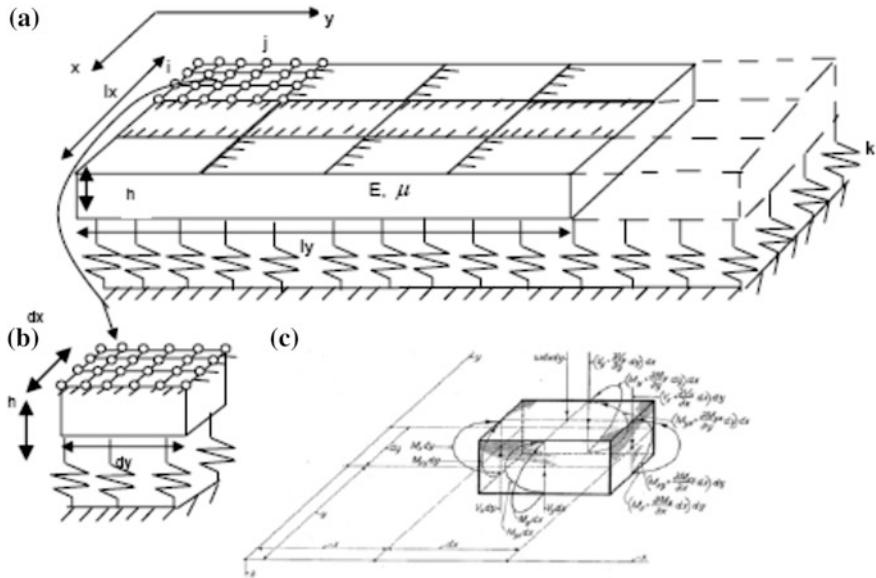


Fig. 2 **a** General concrete road system. **b** Intelligent object (IO) with dx and dy dimensions. **c** Effects on IO under uniform load

numerical solution is done for all these nodes under traffic loadings by the system by intelligent object scanning. Boundary conditions are determined on boundary nodes and the system makes numerical solution according to these boundary conditions. Figure 3 shows the spatial relations of the node.

The general closed equations for the plate on elastic solid are as follows [6]:

$$\Delta \Delta w = (-kw + P)/D \quad (1)$$

$$D = \frac{Eh^3}{12(1 - \mu^2)} \quad (2)$$

$$D \left(\frac{\partial^4 w}{\partial x^4} + 2 \frac{\partial^4 w}{\partial x^2 \partial y^2} + \frac{\partial^4 w}{\partial y^4} \right) = -kw + P \quad (3)$$

w deflection

P applied load per unit area

k coefficient of elastic foundation

D modulus of rigidity of the plate

E Young's modulus of elasticity of the material

h thickness of plate

μ poisson ratio

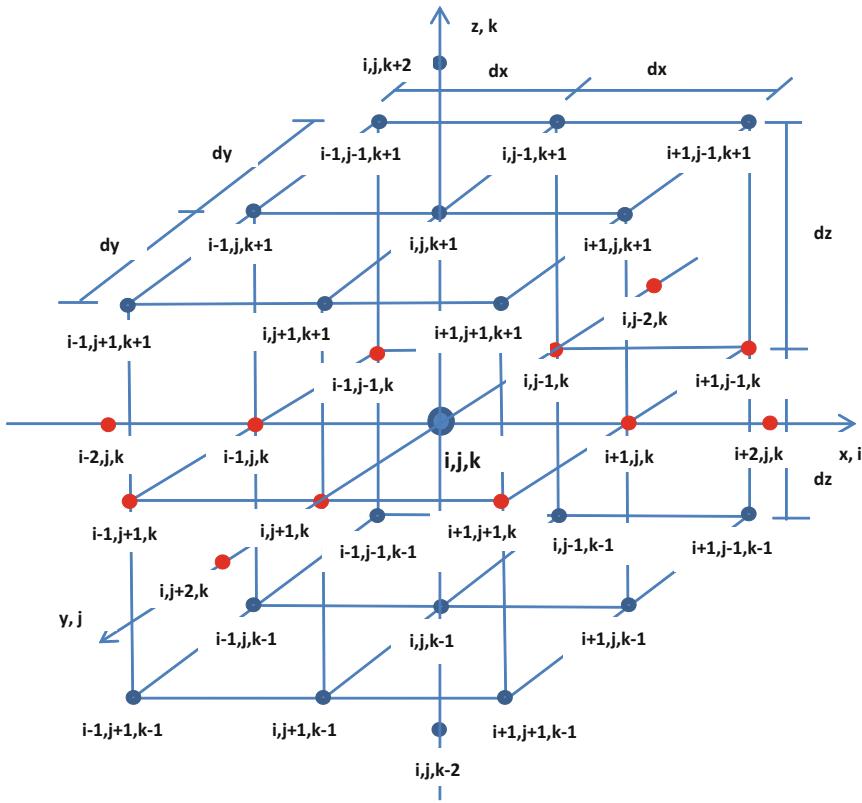


Fig. 3 Spatial relations of the node

Plate on elastic solid;

Let $dx = dy$ (intelligent object length on x and y direction),

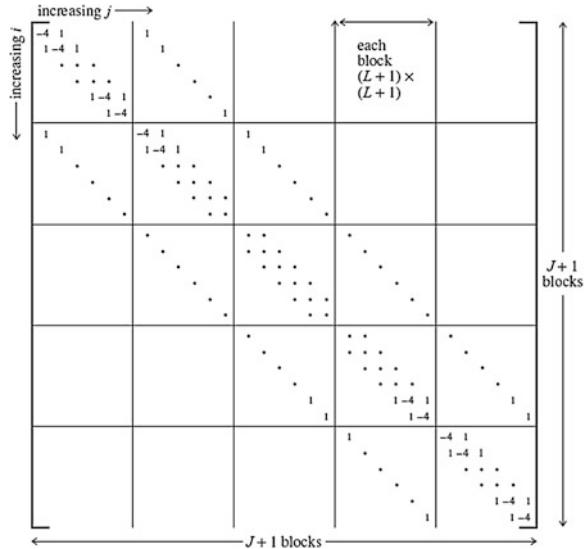
$$[K]\{w\} + k\{w\} = \{P\}/dx^2/D \quad (4)$$

$$[K + k]\{w\} = \{P\}/dx^2/D \quad (5)$$

$$\{w\} = [K + k]^{-1}\{P\}/dx^2/D \quad (6)$$

- [K] Coefficient matrix which is determined by geometric information. In Fig. 4, there is a sample of coefficient matrix form.
- {P} Concentrated load vector (vehicle wheel load)

Fig. 4 Coefficient matrix sample form



Functions (Effect Distribution):

```

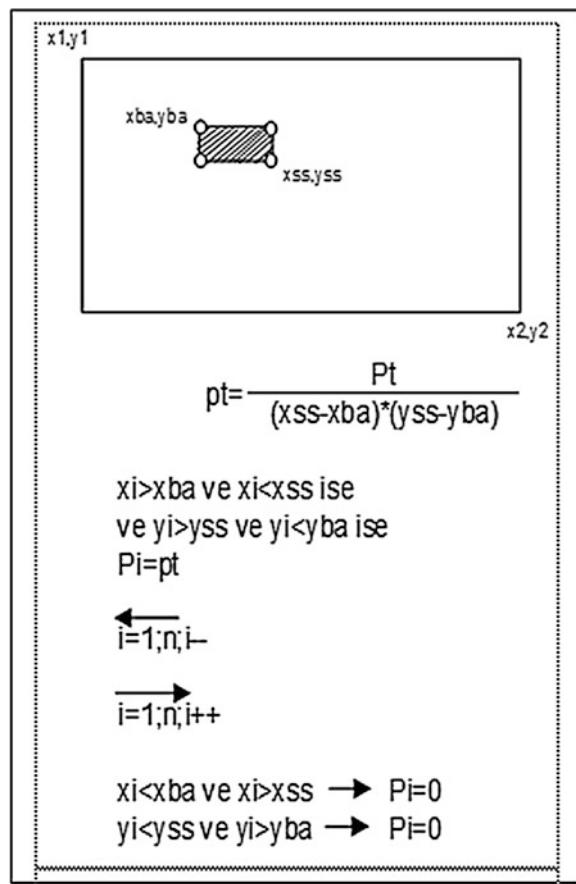
for(i=1; i<=nx; i++)
{
    for(j=1; j<=ny; j++)
    {
        mx[i][j]=kr/(dx*dx)*(2*wc[i][j]-wc[i][j+1]-wc[i][j-1]+(mu*kp*kp)*(2*wc[i][j]-wc[i-1][j]-wc[i+1][j]))
        my[i][j]=kr/(dy*dy)*(kp*kp*(2*wc[i][j]-wc[i-1][j]-wc[i+1][j])+mu*(2*wc[i][j]-wc[i][j+1]-wc[i][j-1]))
        mxy[i][j]=(1-mu)*kr*0.25/(dx*dy)*(wc[i-1][j-1]+wc[i+1][j+1]-(wc[i-1][j+1]+wc[i+1][j-1]))
        qx[i][j]=(mx[i][j+1]-mx[i][j-1])*0.5/dx
        qy[i][j]=(my[i+1][j]-my[i-1][j])*0.5/dy
    }
}

```

$wc[i][j]$, $mx[i][j]$, $my[i][j]$, $mxy[i][j]$, $qx[i][j]$, and $qy[i][j]$ are deflection, moment on x axis, moment on y axis, moment on xy axis, shear stress on x axis, and shear stress on y axis for the nodes on the plate. Figure 5 represents the transformation of axial load to distributed load and plate surface scanning.

It is better to check the numerical solution by comparing it with another numerical solution which is done by Lee [7] using ILLI-SLAB program. In this example, the pertinent input parameters based on the AASHO road test conditions

Fig. 5 Transforming axial load (P_t) to distributed load and plate surface scanning



were: $E = 43.1$ GPa, Load = 80 kN single-axle load, $P = 40$ kN, $\mu = 0.28$, Length of the plate = 4.57 m, Width of the plate = 3.66 m, Loaded area = 279×356 mm², Six slab thickness values of 127, 165, 203, 241, 279, and 318 mm and three k values of 27.1, 40.7, 54.3 kPa/mm were assumed. Lee [7] measured edge stress due to this input parameters. The output is compared with the exact results measured on the site. The wheel load was not directly placed on the slab edge, and the offset distance between the outer face of the wheel and the slab edge was about 330 mm. Fairly good agreement was observed, especially for a k value of 40.7 kPa/mm. In Fig. 5, it is given the output parameters of Lee [7] and the authors using input parameters given above.

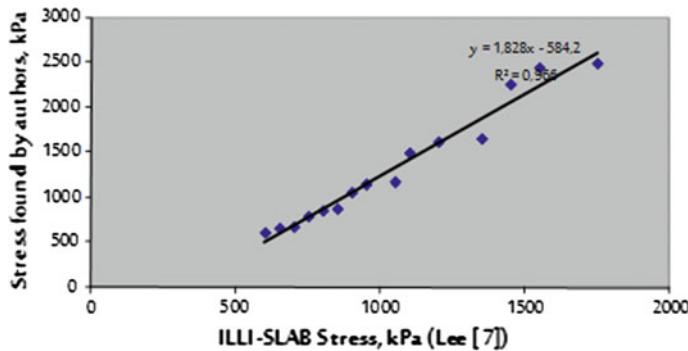


Fig. 6 Stresses found by authors' method and by ILLI-SLAB program

In Fig. 6, 15 different alternatives were compared using ILLI-SLAB program and authors' numerical solution type. Fairly good agreement was observed especially for the slab thicknesses above 203 mm. Optimum concrete road thickness can be found using the reverse computation as explained in this paper.

5 Conclusion

1. Information systems and artificial intelligence technology is a new alternative to manual information processing or computer aided information processing. Concrete road thickness, dilatation places, and reinforcement are determined by the system. This offers the advantage of delivering an optimized result obtained quickly by a deterministic approach.
2. A very large road route system is solved even by computers which have very low capacity. The geometric boundaries are not limited with any shape (no need to be rectangular).
3. The system solution is obtained by dividing general road system into the intelligent objects, which are the subsystem modules and solving these objects. All of the systems are solved by being scanned when this module replaces on the system depending on the system design principles.
4. The system can determine the optimum concrete road design and choose economical concrete road type among jointed plain (JPCP), jointed reinforced (JRCP), and continuously reinforced (CRCP) rigid roads using ISAI technology.

Using information systems and artificial intelligence applied for engineering design stages, optimal numerical solutions can be obtained which are not possible to achieve using traditional methods. In this method, the system gives a precise result using an interior control subsystem. This numerical solution shows that there is a benefit to solve engineering problems using information systems and artificial intelligence technology effectively.

References

1. Khazanovich, L., Selezneva, O., Yu, H.T., Darter, M.I.: Development of rapid solutions for prediction of critical continuously reinforced concrete pavement stress. *Transp. Res. Rec. Design Rehabil. Pavements* **1778**, 64–72 (2001)
2. Bayrak, M.B., Ceylan, H.: Neural network-based approach for the analysis of rigid pavement systems using deflection data. *Transp. Res. Rec.* **2068**, 61–70 (2008)
3. Anastasios, M.I., Don, R.A., Michael, I.H., Craig, M.D.: Application of artificial neural networks to concrete pavement joint evaluation. *Transp. Res. Rec.* **1540**, 56–64 (1996)
4. Kasthurirangan, G., Marshall, R.T., Anshu, M.: Rapid finite-element based airport pavement moduli solutions using neural networks. *Int. J. Eng. Math. Sci.* **3**, 63–71 (2007)
5. Ceylan, H., Tutumluer, E., Barenberg, E.J.: A structural model for the rapid analysis of concrete pavement systems. *Mid-Continent Transportation Research Symposium*, Ames, Iowa, Aug 2003
6. Berkay, I.: *Plak Teorisi ve Uygulamaları*. Yıldız Teknik Üniversitesi Muh.Fak.Insaat Muh. Bolumu, İstanbul, pp. 237 (1992)
7. Lee, Y.H.: TKUPAV: stress analysis and thickness design program for rigid pavements. *J. Transp. Eng.* **125**, 338–346 (1999)

Multi-hop Secured Video Transmission Using Dual-channel Dual-path Routing Against Packet Copy Attack

Vaibhav Mishra, Saurabh Maheshwari and Kavita Choudhary

Abstract The video data that are used in the multimedia applications are not completely secure because of multiple unsecured hops present between the sender and the receiver. Any intermediate node may behave malicious and copy of the data packets illegally. This paper introduces a novel method to send video data through dual channel using dual data path. The frames' pixels are scrambled. The video frames are divided into two frame streams. At the receiver side, video is reconstructed and played for a limited time period. As soon as small chunk of merged video is played, it is deleted from video buffer. The approach has been tried to formalize, and initial simulation has been done over MATLAB. Preliminary results are optimistic, and a refined approach may lead to a formal designing of network layer routing protocol.

Keywords Multi-path · Multi-hop · Dual-channel · Scrambling · Video transmission · Video encryption · Routing · Copy attack · Security

1 Introduction

Unsecured network attracts more intruders or hackers to gain information in an illegal way. The presence of the multi-path routing is the hurdle for them. The data need to be transmitted through multiple paths. The network is very vast, and the data are moving from one place to other. The main objective is to remove the problem of

V. Mishra (✉) · K. Choudhary
JIET Jodhpur, Jodhpur, Rajasthan 342802, India
e-mail: vaibhav.mishra071@gmail.com

K. Choudhary
e-mail: kavity_26@yahoo.co.in

S. Maheshwari
GWEC Ajmer, Makhupura Ajmer, Rajasthan 305002, India
e-mail: saurabh.maheshwari.in@ieee.org

data theft which is done by the various malicious nodes which are present between the source and destination. The secure transmission of the video is again a challenge in the network. The objective is a secure transmission of the video from source to destination. The data present in the network are in such a form so that if the malicious attacker gets the data, he cannot be able to construct an original form of video. This provides a high-level security so that the video data are protected from the intruders. In this work, we propose to divide the data into the two streams and transmitting it from two different paths rendering it difficult to predict the original content of the data [1]. The presence of the multi-path and data in multiple channels ensures novel way of source to destination data transfer. Multi-path routing has various applications such as video conferencing, IP network multi-hop routing, and Internet protocol television (IPTV) security and secure virtual paths. IPTV user pays for the services that he want to watch at home. In both the paths, the data packets are in random order that no one can predict the actual position of the data packet, i.e., the packet is present in the left channel or right channel. The packets are in scrambled form so that originality of the data is maintained, and if any person copies the data packet in between the sender and the receiver by attacking, he cannot be able to construct a original video because the data are in the encrypted form [2]. There are various benefits of the multi-channel such as security of the data packets, fast transfer of data from sender to receiver, and continuous streaming of data at receiver side. As per the conventional view, the data are send from source to destination. These data move from one side to other side by using the same path. These data that are traveling from source to destination move from different nodes; sometimes, these nodes act maliciously. The probability of theft of data increases. The important data get leaked by the malicious nodes which may lead to the troublesome problem of piracy of data. This paper is structured in the various sections. Section 2 discusses the various video encryption methods; real-time constraints, which affect the quality of the video on the network; and various solutions which are concerned with the successful transfer of video from sender to receiver. Section 3 focuses on the approach of sending the data packets through proposed dual channel dual path. Section 4 presents the results and comparisons of the different approaches which show the resistance against malicious nodes on data packets. Section 5 contains the conclusion and future aspects of the work.

2 Literature Survey

2.1 Security for IPTV

IPTV is an emerging technology which uses the traditional television services like cable TV. The services of IPTV need communication of high-definition video. Being a new area, there are many chances of making illegal copy and data. Watermarking, fingerprinting, scrambling, and encryption techniques are applied to provide the

security to the content of the IPTV. Some of the general security measures which are taken are authentication, distribution of keys, authorization, etc. [1]. In multimedia applications, the video data are secured by applying the video encryption methods to the data. To avoid the predictability of the original video content, different video encryption techniques are applied. Various encryption techniques such as scalable frame scrambling, selective encryption, and secret key cryptography are boon in providing the security to the data. In the scalable frame scrambling technique, the ordering of the frames is broken by internal scrambling of the frames. Different levels of security are applied on these frames so that the security level increases which prevents the data theft. On the next higher level of security, the internal scrambling of the pixels changes the original video which is worthless for the intruder [3]. In compressed bit stream methods, the bit streams are compressed with the help of the DES key and this DES key is encrypted with the help of RSA key. The security of the multimedia applications is an important concern [4]. The method of taking samples from the video stream for encryption instead of encrypting whole video stream is called perceptibility [5]. Video stream is divided into the different streams, and each stream is send by the different path [6].

2.2 Real-time Constraints

When the scalable video is send from the sender to the receiver, there are lot of real-time constraints present in single path or multi-path. The problem of statistical delay bound occurs with the scalable video. Reliable delivery of the video is also major concern in the video surveillance applications because the size of the video changes and energy constraint of the nodes is also an issue. These factors affect the reliability of the network therefore video data should be sent through different paths [7, 8].

2.3 Multi-channel Concept

Multi-channel multi-hop communication has some problems related to the interference which decreases the quality of the video because the noise is introduced into the main data signal; therefore, a new CSMA/CA-based scheme is introduced which reduces the inter-path and intra-path interference and AODV is used for assigning the channels [9]. In Hierarchical routing, the clusters are formed and beacon messages are used in the formation of the clusters [3]. Reliable delivery of the video from multiple paths is a major concern in the video surveillance applications because the size of the video changes and energy constraint of the nodes is also an issue [7]. Video streaming is a very important concept in the multimedia applications. In multiple description of video transmission, the video is divided into the different streams and each stream is

send by the different path [6]. The video that is send over the network is evaluated based on the various parameters such as PSNR, delay, and throughput, because there are many factors which affect the quality of the video in the network; therefore, the opportunistic scheme is made in the quality of the video [10]. Conventionally a video data is transferred through a single path but the opportunistic scheme, first the nearest hop is found and then video data starts moving from the sender to the receiver [4].

3 Proposed Method

The illegal users and the malicious subscribers who have wrong intentions copy this video data and upload them over Internet that bring loss of millions to the digital TV world and their service qualities. This copy generally takes place either at the receiver device or at the intermediate routing nodes in the network. A security mechanism is required between the video senders and the video receivers. Both should run a secure protocol which not only encrypt the video data but also avoid illegal copying and playing. Generally, videos have RGB-layered frames. Here, the flowchart is given which shows the step-by-step working of the proposed algorithm. According to the proposed algorithm, first the graphical architecture between the source and destination is formed, and after that, the positions of the nodes, i.e., (x, y) coordinates of the nodes, are fixed. Initial routing tables are created which contain the address of the one hop neighbors. Sender looks for the id of the receiver, and then, handshaking between the sender and receiver is done. Acknowledgment signal is send by the receiver to sender on successful handshaking. Video information is shared between the sender and destination. Channel is divided into two data streams. Random paths will be created between the source and destination. Bit streams are collected separately at the receiver side, and video is played for a limited time period only and nothing will be stored into the buffer the steps of proposed method is shown in Fig. 1.

4 Simulation Results

The proposed algorithm has been simulated using MATLAB as simulation tool. Generally, the network simulators are used for routing protocol simulations and validations. Here, we are utilizing the MATLAB as network simulation tool for video communication due to ease of video processing in MATLAB. The nodes and packets are dealt as structures with the packet fields, and node parameters are structure variables. The algorithm has been tested for variety of values of simulation parameters. We will discuss the simulation results on the following values of the simulation parameters. The plots of original frame and the R, G, and B layers are shown in Fig. 2. Overlapping of these layers makes the colored video frames. These individual layers

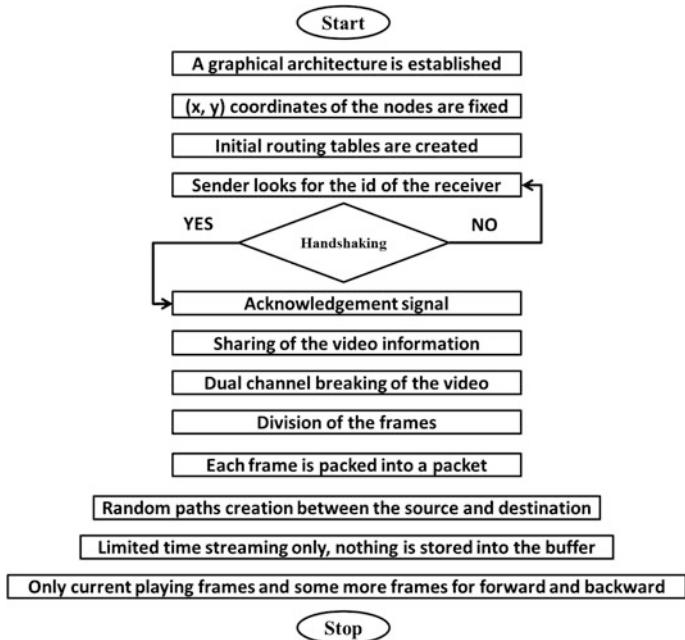


Fig. 1 Flowchart of proposed algorithm



Fig. 2 Original frame, R, G, and B frames

will be embedded in a packet to be transferred to the receiver. Receiver after getting them overlaps them in correct order to make the playable frame (Table 1 represents various simulation parameters used to test the algorithm).

The network architecture for simulation is given in Fig. 3. The figures below show the data transmission path from sender to receiver through intermediate nodes. The first node in the path is the sender. Snapshots of the running transmission of a certain packet through path 1 and path 2 are shown in Fig. 4. In simulation, the nodes in blue connected with black path lines depict the path 1 nodes and left channel data flow, respectively. Different packets reach to the receiver. Receiver collects them in receiver buffer. This path is constant for entire transmission (Fig. 3).

Table 1 Simulation parameters

Parameters	Symbol	Value
Total nodes	N	100
Pre-shared secret key	K	12,345,678
Video frames	F	200
Video height	H	240 pixels
Video width	W	320 pixels
Video format	Fmt	'RGB24'
Frame rate	R	30 FPS
Video duration	T	~ 7 s
Bit size per pixel	B	24 bits or bits per layer
Number of nodes in path 1	N1	30
Number of nodes in path 2	N2	30
Malicious nodes in path 1	M1	10
Malicious nodes in path 2	M2	10
Channel capacity	C	2,048 Kbps
Data transfer rate	Tr	1,024 Kbps
Packet size	$P = (H * W * 8)/100$	614 Kb
Time to play chunk	t	5 s

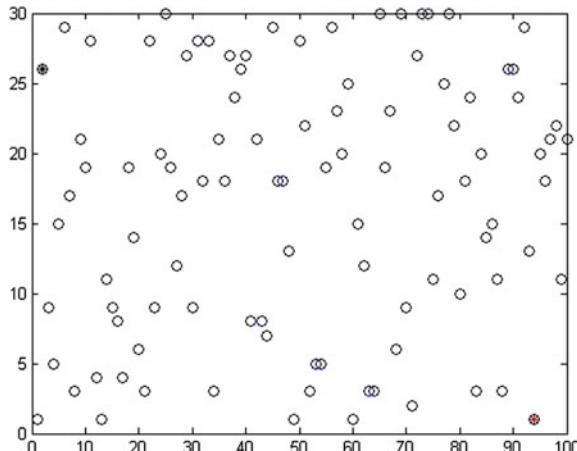
**Fig. 3** Network architecture

Figure 5 displays the video player, playing the chunks of the video. The algorithm has been designed such that there is no lag between the chunks of the video. It appears to the viewer that the video is continuously playing, but in fact, it is in small chunks which are regularly deleted from the video memory after playback. To restrain packet copy attack, data are divided into two paths using dual-channel

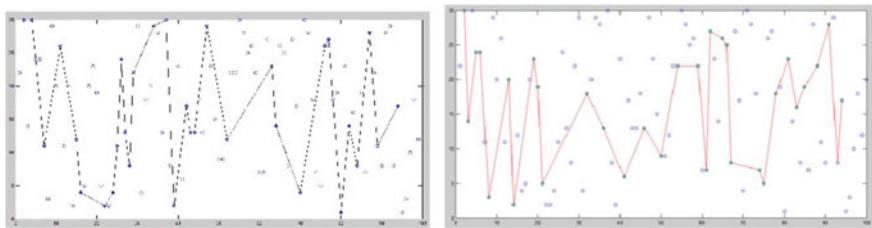


Fig. 4 Data received by receiver through path 1 and 2

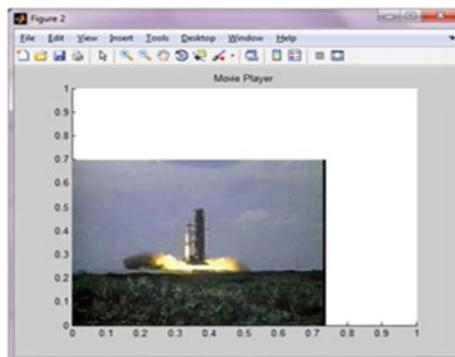


Fig. 5 Video player at receiver playing in small chunks

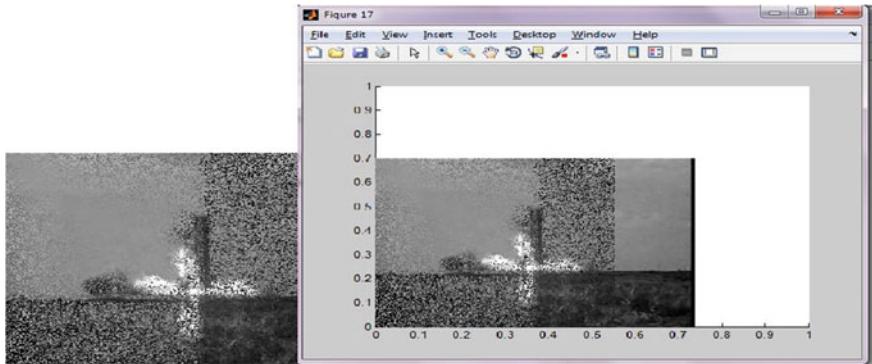
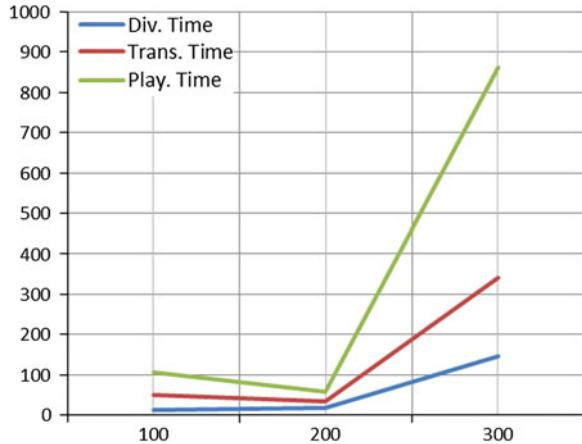


Fig. 6 Scrambled video frame layer and malicious node external storage movie frames played

concept. Now, all the packets do not reach to a single node. Moreover, the packets copied in external storage are in scrambled format as shown in Fig. 6. We also played the video data collected by malicious nodes in their external storages. The

Fig. 7 Comparison of division time, transmission time, and playback time with respect to number of frames for DCDP approach



video played is shown in Fig. 6. Due to dual-channel division, scrambling the quality of the video is poor (Fig. 6).

The actual playback at the receiver is shown in Fig. 5. This is not a continuous playback. It may appear to the viewer of the video that the video is continuously playing but actually only a chunk of video is being merged from the incoming dual-channel packets and played.

4.1 Effective Runtime Storage

The total runtime storage needed for the video transmission at the sender, receiver, and intermediate nodes can be calculated. Size of complete video is $W * H * 24 * F$ bits. The receiver needs to play the video for specific time t only. So the maximum buffer needed to be allocated for playback is $t * F * W * H * 24$ bits. Intermediate nodes only forward the data packet to next hop node in the paths (Figs. 5 and 7).

4.2 Comparative Results

Comparison of division time, transmission time, and playback time with respect to number of frames for Dual channel Dual path approach is given in Fig 7. Comparison of single-channel dual-path (SCDP), dual-channel single-path (DCSP), and single-channel single-path (SCSP) approaches show satisfactory result. In DCSP approach, the video data are divided into two channels, but the data transmission is done through single path. The transmission time comparison can be done for both the strategies. It can be seen from the Fig. 8 that the time taken by DCDP algorithm

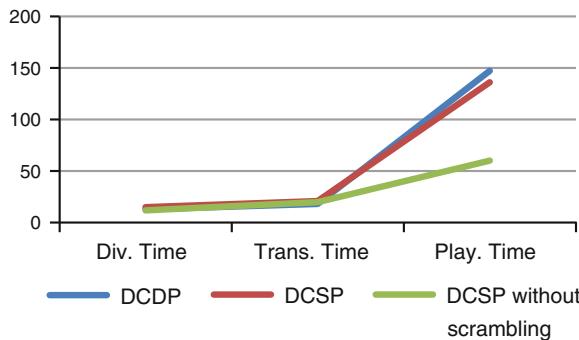


Fig. 8 Comparison of DCDP, DCSP, and DCSP without scrambling approaches with respect to division time, transmission time, and playback time

is almost half of the time taken by DCSP approach. This is due to the fact that in the DCDP approach, the data are divided into two almost equal data channels and transmitted parallel through two paths, so overall time needed becomes half.

The time taken for division is maximum for DCSP algorithm, whereas the transmission time is minimum in DCDP as shown in Fig 8. This is because in DCDP, multiple paths are used to transmit the dual-channel data, whereas in DCSP, single path is used so concurrency is reduced and time increases. Playback time is smallest in DCSP without scrambling, but this is not attack resistant since all packets move from single path, and without scrambling, the copied packets can be visually interpreted. In DCSP method, the scrambling concept is not present and the intermediate node gets all the packets since single path is followed and the video will be played in correct format therefore the attack is successful. SCSP approach is followed for normal data transmission done in IP network where the complete data transmitted through a single constant path. Complete set of packets are available at all the nodes of the network so attacker can copy it from any node. The main disadvantage of this approach over DCDP is that the overall transmission time will be almost double since all the packets move through single path and reach at receiver in serial communication mode. SCDP approach also shares the disadvantages of SCSP and DCSP approaches. It also adds another disadvantage due to single-channel transmission through two paths. The network now has duplicate packets. Duplicity is a big problem since it needs all the network resources in twice amount. So this is not a feasible way for communication. Peak signal-to-noise ratio (PSNR) is a quality metric to measure the quality of signals after processing in comparison with the original signals. PSNR [8] between the original video frames and the merged frames at the receiver is calculated from simulation. The PSNR is found infinite, which suggests that there is no visible difference between the original and the transmitted frame. PSNR between the original frame and the frames generated by malicious nodes is less than 15 dB.

5 Conclusion and Future Work

The work described in this thesis is a novel strategy to route the video packets from sender to receiver using multi-channel and multi-path. We have shown how the proposed DCDP method is best among current strategies such as DCSP, SCSP, and SCDP. We also showed how the video is non-watchable at the malicious nodes which are trying to illegally copy video packets. The PSNR metric calculated between the original and the received video frames is sufficiently high for proposed method, and it is sufficiently low for video frames copied by malicious nodes. Comparison of methods in similar domain has been done for dual-channel division time, transmission time, and playback time. Most important advantage of the algorithm is that there is no playback lag at the receiver while it is playing the video in the form of chunks. The approach can be well utilized for transmission of self-destroying video messages used by military and intelligence agencies. If any application needs to scroll the video to previous playback chunk, then it is not possible in current method. That chunk is to be again created and then played.

References

1. Hwang, S.O.: Content and service protection for IPTV. *IEEE Trans. Broadcast.* **55**(2), 425–436 (2009)
2. Khalek, A.A., Dawy, Z.: Optimized resource allocation for video transmission over multihop networks with end-to-end statistical delay bound. In: 2011 19th International Conference on Software, Telecommunications and Computer Networks (SoftCOM), pp. 1–6, 15–17 Sept 2011 (2011)
3. Rajpurohit, J., Khunteta, A.: A scalable frame scrambling algorithm for video encryption. In: 2013 IEEE Conference on Information and Communication Technologies (ICT), pp. 981–985, 11–12 April 2013
4. Wu, D., Ci, S., Wang, H., Katsaggelos, A.K.: Application-centric routing for video streaming over MultiHop wireless networks. *IEEE Trans. Circ. Syst. Video Technol.* **20**(12), 1721–1734 (2010)
5. Jeffrey, M., Park, S., Lee, K., Adams, G., Savage, S.: Content security for IPTV. *IEEE Commun. Mag.* **46**(11), 138–146 (2008)
6. Rong, B., Qian, Y., Lu, K., Hu, R.Q., Kadoch, M.: Multipath routing over wireless mesh networks for multiple description video transmission. *IEEE J. Sel. Areas Commun.* **28**(3), 321–331 (2010)
7. Do Rosario, D., Costa, R., Paraense, H., Machado, K., Cerqueira, E., Braun, T.: A smart multi-hop hierarchical routing protocol for efficient video communication over wireless multimedia sensor networks. In: 2012 IEEE International Conference on Communications (ICC), pp. 6530–6534, 10–15 June 2012
8. Maheshwari, S., Gunjan, R., Laxmi, V., Gaur, M.S.: Robust multi-modal watermarking using visually encrypted watermark. In: The 19th International Conference on Systems, Signals and Image Processing, IWSSIP 2012, IEEE, pp. 72–75, Vienna, Austria 11–13 April 2012

9. Qaisar, S.B., Radha, H.: Multipath multi-stream distributed reliable video delivery in Wireless Sensor Networks. In: 43rd Annual Conference on Information Sciences and Systems, 2009, CISS 2009, pp. 207–212, 18–20 March 2009
10. Xu, Y., Deng, J.D., Nowostawski, M.: Optimizing routing in multi-hop wireless networks using analytical capacity estimation: a study on video streaming. In: 2012 IEEE 14th International Conference on High Performance Computing and Communication and 2012 IEEE 9th International conference on Embedded Software and Systems (HPCCICESS), pp. 748–755, 25–27 June 2012

Static Analysis: A Survey of Techniques and Tools

Anjana Gosain and Ganga Sharma

Abstract Static program analysis has shown tremendous surge from basic compiler optimization technique to becoming a major role player in correctness and verification of software. Because of its rich theoretical background, static analysis is in a good position to help produce quality software. This paper provides an overview of the existing static analysis techniques and tools. Further, it gives a critique of static analysis approach over six attributes, namely precision, efficiency, coverage, modularity, scalability, and automation.

Keywords Static analysis · Data flow analysis · Abstract interpretation · Constraint solving · Symbolic execution · Theorem proving

1 Introduction

Static analysis is the process of automatically analyzing the behavior of computer programs without executing it [1–3]. Dynamic analysis, on the other hand, analyses programs by executing an instrumented program and generating some form of trace [4]. Static analysis techniques have remained the most popular choice for analysis of computer programs because of the fact that it is very simple and fast [5]. Dynamic analysis has existed in parallel with static analysis providing a complementary alternative to static analysis [4, 6, 7]. This paper is an attempt to provide an overview of static analysis and its associated techniques and tools. The remaining paper is organized as follows: Sect. 2 and its subsections give brief introduction of static program analysis, its taxonomy, and associated techniques and tools.

A. Gosain (✉) · G. Sharma

USICT, Guru Gobind Singh Indraprastha University, New Delhi 110078, India
e-mail: anjana_gosain@hotmail.com

G. Sharma

e-mail: ganga.negi@gmail.com

Section 3 gives a critique of static analysis, namely precision, efficiency, coverage, modularity, scalability, and automation. Finally, Sect. 4 concludes the paper.

2 Static Analysis

Static analysis means automatic methods which can reason about run-time properties of program without actually executing it [2, 8]. Static analysis has been used to detect errors which might lead to premature termination or ill-defined results of the program [9]. Following is a non-exhaustive list of errors that can be addressed using static analysis [3, 9, 10]:

- array out of bound, i.e., accessing an element of array beyond its size.
- null pointer dereference, i.e., a pointer with NULL value is used as though it contained a valid memory address.
- memory leaks, i.e., a program fails to return memory taken for temporary use.
- buffer underflow/overflow, i.e., size of a data value/memory used for holding is not taken care of.
- access to uninitialized variables/pointers, i.e., use of variables/pointers without assigning them values/memory.
- invalid arithmetic operation, e.g., division by zero.
- non-terminating loops, e.g., exit condition of the loop does not evaluate to false.
- Non-terminating calls, i.e., the control flow of a program never returns from a function call.

Most of the aforementioned errors can be detected by software testing. Although testing still remains the major validation activity in software development process [8], the promising nature of program analysis approaches can help software developers in producing correct and reliable software. The main advantages and disadvantages of static analysis over testing are summarized in Table 1.

One of the key ideas behind static analysis is abstraction. Abstraction refers to transformation of a program, called concrete program, into another program that still has some key properties of the concrete program, but is much simpler, and therefore easier to analyze [5]. Over-approximation results when there are more behaviors in the abstraction than in the concrete system [5]. Under-approximation, on the other hand, deals with fewer behaviors than in the concrete system.

Static analysis can be sound or unsound. Soundness guarantees that the information computed by the analysis holds for all program executions, whereas unsound analysis does not [11]. Static analysis can be made path, flow, and context sensitive by making it to distinguish between paths, order of execution of statements, and method call (call site), respectively [12]. Precision of an analysis approach can be measured by classifying the analysis results into three categories [13]: false positives, i.e., nonexistent bugs are reported as potential bugs; false negatives, i.e., bugs that are undiscovered; and true positives, i.e., true and discovered bug. Efficiency is related to computational complexity or cost pertaining to

Table 1 Comparison of static analysis and testing

Static analysis	Testing
Can be applied without executing the code	Can be applied only by executing the code
Can be applied early in the development process	Is applied late in the development process
Results do not depend on inputs	Results depend on inputs
Results can be generalized for future executions	Results cannot be generalized for future executions
Less costly	Very costly
Very fast process	Slow process
False-positive rate is very high	False-positive rate is very less
Approximations are used	Exact results are used
Cannot be used for functional correctness of program	Can be used for functional correctness of programs

space and time requirements of the algorithms used in the analysis [5]. Precision and efficiency are related to each other, i.e., a precise analysis is more costly and vice versa.

2.1 Static Analysis Techniques

2.1.1 Syntactic Pattern Matching

This technique is based on the syntactic analysis of a program by a parser. The parser takes as input the program source code and outputs a data structure called abstract syntax tree [1]. When used for bug finding, this technique works by defining a set of program constructs that are potentially dangerous or invalid and then searching the input program's abstract syntax tree for instances of any of these constructs. Syntactic pattern matching is the fastest and easiest technique for static analysis. But it provides little confidence in program correctness and can result in a high number of false alarms [14].

2.1.2 Data Flow Analysis

It is the most popular static analysis technique. The general idea was given in [15–17]. It proceeds by constructing a graph-based representation of the program, called control flow graph, and writing data flow equations for each node of the graph. These equations are then repeatedly solved to calculate output from input at each node locally until the system of equations stabilizes or reaches a fixed point. The major data flow analyses used are reaching definitions (i.e., most recent assignment to a variable), live variable analysis (i.e., elimination of unused assignments), available

expression analysis (i.e., elimination of redundant arithmetic expressions), and very busy expression analysis (i.e., hoisting of arithmetic expressions computed on multiple paths) [2]. At each source code location, data flow analysis records a set of facts about all variables currently in scope. In addition to the set of facts to be tracked, the analysis defines a “kills” set and a “gens” set for each block. The “kills” set describes the set of facts that are invalidated by execution of the statements in the block, and the “gens” set describes the set of facts that are generated by the execution of the statements in the block. To analyze a program, the analysis tool begins with an initial set of facts and updates it according to the “kills” set and “gens” set for each statement of the program in sequence. Although mostly used in compiler optimization [16, 17], data flow analysis has been an integral part of most static analysis tools [18–21].

2.1.3 Abstract Interpretation

This technique was formalized by Cousot and Cousot [22]. It is a theory of semantics approximation of a program based on monotonic functions over ordered sets, especially lattices [23]. The main idea behind this theory is as follows: A concrete program, its concrete domain, and semantics operations are replaced by an approximate program in some abstract domain and abstract semantic operations. Let L be an ordered set, called a concrete set, and let L' be another ordered set, called an abstract set. A function α is called an abstraction function if it maps an element x in the concrete set L to an element $\alpha(x)$ in the abstract set L' . That is, element $\alpha(x)$ in L' is the abstraction of x in L . A function γ is called a concretization function if it maps an element x' in the abstract set L' to an element $\gamma(x')$ in the concrete set L . That is, element $\gamma(x')$ in L is a concretization of x' in L' . Let $L1, L2, L'1$, and $L'2$ be ordered sets. The concrete semantics f is a monotonic function from $L1$ to $L2$. A function f' from $L'1$ to $L'2$ is said to be a valid abstraction off if for all x' in $L'1$, $(f \circ \gamma)(x') \leq (\gamma \circ f')(x')$. The primary challenge to applying abstract interpretation is the design of the abstract domain of reasoning [24]. If the domain is too abstract, then precision is lost, resulting in valid programs being rejected. If the domain is too concrete, then analysis becomes computationally infeasible. Yet, it is a powerful technique because it can be used to verify program correctness properties and prove absence of errors [25, 26].

2.1.4 Constraint-Based Analysis

A constraint-based analysis traverses a program, emitting and solving constraints describing properties of a program [27, 28]. This technique works in two steps. The first step, called constraint generation, produces constraints from the program text, which describe the information or behavior desired from the program. The second step is constraint resolution which involves solving the constraint by computing the desired information. Static information is then extracted from these solutions.

One of the key features of this technique is that algorithms used for constraint resolution can be written independently of the eventual constraint system used [28].

2.2 Miscellaneous Techniques

2.2.1 Symbolic Execution

In this method, instead of supplying normal inputs to the program, one uses symbols representing arbitrary values [29]. The execution proceeds as in a normal execution except that the values are now symbolic formulas over input values. As a result, the output values computed by a program are expressed as a function of the input symbolic values. The state of a symbolically executed program includes the symbolic values of program variables, a path condition (PC) and a program counter. The path condition is quantifier-free Boolean formula over the symbolic inputs; it accumulates constraints which the inputs must satisfy in order for an execution to follow the particular associated path. A symbolic execution tree characterizes the execution paths followed during the symbolic execution of a program. The tree nodes represent program states, and they are connected by program transitions. Symbolic execution is the underlying technique of several successful bug-finding tools like PREfix [18] and CodeSonar [19].

2.2.2 Theorem Proving

Theorem Proving is based on the deductive logic proposed by Floyd and Hoare [30, 31]. A program statement S is represented as a triple $\{p\}S\{q\}$, where p (precondition) and q (post-condition) are logical formulas over program states. This triple is valid iff for a state t satisfying formula p , executing S on t yields a state t' which satisfies q . Various inference rules are then used to verify system states. One of the most famous theorem provers is Simplify [32] which has been used in tools like ESC/Java [21]. Tiwari and Gulwani [33] have used theorem proving for a new technique called logical interpretation, i.e., abstract interpretation over logical lattices. This technique tries to embed theorem proving technology in the form of engines of proof into the static analysis tools based on abstract interpretation. Then, it is used for checking as well as generating program invariants.

2.3 Static Analysis Tools

There are numerous commercial as well as open-source static analysis tools available. A summary of these tools is given in Table 2. We have categorized these tools in the following categories based on the way they are used:

Table 2 Summary of static analysis tools

Tools	Static checking	Error detection/bug finding	Verification	Type inference
Lint family [34–36]	✓			
JLint [37]	✓			
FindBugs [20]	✓	✓		
PMD [39]	✓	✓		
Coverity Prevent [41]		✓		
KlockWork K7 [40]		✓		
ASTREE [25]		✓	✓	
CGS [10]			✓	
Polyspace Verifier [26]			✓	
TVLA [42]			✓	
BANE [44]				✓
BANSHEE [45]				✓
CQual [43]				✓
PREfix [18]		✓		
CodeSonar [19]		✓		
ESC/Java [38]	✓		✓	
ESP [21]		✓	✓	

- Static checking [20, 34–39]: The aim here is to find common coding errors.
- Error detection/bug finding [18, 19, 40, 41]: The aim of the tools in this category is to find bugs and report them.
- Software verification [10, 25, 26, 42]: The tools in this category guarantee the absence of errors by providing proofs.
- Type qualifier inference [43–45]: The tools in this category specify as well as check program properties.

The first static analysis tool was Lint [34]. It was a simple Unix utility command which accepted multiple files and library specifications and checked them for inconsistencies like unused variables and functions, unreachable code, non-portable character, and pointer alignment. But Lint could not identify defects that can cause run-time problems. A series of shallow static analysis tools like FlexeLint/PCLint [35] and Splint [36] then followed. These tools relied on the syntactic information, and the analysis is mainly path or context sensitive. Splint additionally checks for security vulnerabilities. These tools mainly worked for C/C++. JLint [37] is the Lint family tool to check Java code for inconsistencies, bugs, and synchronization problems. Other tools for Java which use syntactic pattern matching include FindBugs [20] and PMD [39]. FindBugs also uses a simple, intra-procedural data flow analysis to check for null pointer dereferences. PMD is an open-source, rule-based, static analyzer that analyzes Java source code based on the evaluative rules

that have been enabled during a given execution. The tool comes with a default set of rules which can be used to unearth common development mistakes such as having empty try-catch blocks, variables that are never used, and objects that are unnecessary.

ESP [21] uses path-sensitive data flow analysis for error detection. It uses a combination of scalable alias analysis and property simulation to verify that large C code bases obey temporal safety properties. Symbolic execution is the underlying technique of the successful bug-finding tool PREfix for C and C++ programs [18]. For each procedure, PREfix synthesizes a set of execution paths, called a model. Models are used to reason about calls, and then, fixpoints of models are approximated iteratively for recursive and mutually recursive calls. Polyspace Verifier [26] uses sound analysis based on abstract interpretation which can prove absence of run-time errors related to arithmetic operations. Prevent [41] uses path simulation techniques to prune the infeasible paths and helps in curbing false alarms. CodeSonar [19] uses symbolic execution engine to explore program paths, reasoning about program variables and how they relate. It then uses data flow analysis to prune infeasible program paths from being explored. Only Klockwork K7 [40] supports software metrics in this category apart from other common facilities. CGS [10] is a precise and scalable static analyzer which uses abstract interpretation to check run-time errors in embedded systems and has been used extensively on real NASA space missions.

ASTREE [25] has similar foundational theory as CGS but also is sound. It is used for checking run-time errors as well as assertion violations and has been used in safety-critical systems. Three Valued Logic Analysis Engine (TVLA) [42] is a system that can automatically generate abstract interpretation analysis algorithms from program semantics. It has been successfully used in performing shape analysis of heap allocated data [46]. The most popular static analysis tool which uses theorem proving is ESC/Java [38]. To use ESC/Java, the programmer adds pre-conditions, post conditions, and loop invariants to source code in the form of special comments called annotations. It then uses Simplify theorem prover [32] to verify that the program matches the specifications.

CQual [43] is a constraint-based analysis tool used for type inference. It traverses the program's abstract syntax tree and generates a series of constraints that capture the relations between type qualifiers. If the constraints have no solution, then there is a type qualifier inconsistency, indicating a potential bug. BANE [44] is a framework which helps in developing constraint-based program analyses. It frees the user from writing the constraint solver which is the most difficult task in constraint-based analysis. The user only writes code to generate constraints from program texts. Banshee [45] is a scalable version of BANE. It inherits a lot of features of BANE and has added the support for incremental analysis using backtracking.

3 A Critique of Static Analysis

3.1 Precision

Static analysis can be imprecise because it makes assumptions or approximations about the run-time behavior of a program. Therefore, it suffers a lot from false positives. One way to avoid false positives is to filter the analysis results, removing errors which are highly unlikely. For example, in PREfix [18], the user of the tool can control the maximum number of execution paths that will be generated for a function. Another way to avoid false positives is to prune the paths which may be infeasible. For example, Coverity Prevent [40] uses SAT solvers for this kind of false path pruning. But filtering or pruning can introduce false negatives. Therefore, use of filtering/pruning should be done with caution.

3.2 Efficiency

Efficiency is an important attribute as it directly links to the cost of the analysis. Efficiency is also related to precision. Static analysis trades off precision over efficiency; that is, in order to provide analysis results faster, it uses approximation or abstraction mechanisms which lead to less precise results. But there exist techniques which can make static analysis precise by making static analysis flow, path, or context sensitive [12]. But this precision comes at the cost of longer time and greater resource need.

3.3 Coverage

It implies the total number of execution paths analyzed [5, 11]. Obviously, the execution paths analyzed must be valid ones. By its very nature, static analysis provides high coverage as it does not depend on specific input stimuli. Therefore, as compared to its dynamic counterparts, static analysis provides high coverage of code for analysis purpose.

3.4 Scalability

In general, a system is said to be scalable if it remains suitably efficient and practical when applied to situations in the large [5]. This is required when one applies analysis approaches to industrial-size software. Incorporating scalability into a static analysis

is tedious as it operates on some kind of model of the system. Most static analysis tools employing abstract interpretation [10, 25, 26] have been developed to pursue scalability. ESP [21] is made scalable using property simulation.

3.5 Modularity

A program is analyzed by analyzing its parts (such as components, modules, and classes) separately, and then, analysis results are composed together to get the required information on the whole program [24]. The advantage of using this approach is that precision, efficiency, and scalability can be greatly enhanced. Work has been done on modular static analysis by Cousot and Cousot [47] to enhance scalability. Dillig [48] describes a fully modular, summary-based pointer analysis that can systematically perform strong updates to abstract memory locations reachable through function arguments. ESC/Java [38] is a static analysis tool which supports modular analysis.

3.6 Automation

It implies the extent to which the analyzer analyses the code by itself without requiring human intervention [12]. Static analysis is a fully automated technique and thus is the most popular choice for analysis of source code [5].

4 Conclusion

This paper studies the foundational techniques of static program analysis. It provides a list of static analysis tools with a brief overview of these tools. It further provides a critique of static analysis over six attributes, namely precision, efficiency, coverage, scalability, modularity, and automation. Static analysis is the most widely used technique for program analysis because of its high efficiency, coverage, and automation, but suffers from high false-positive rates. Still, static analysis techniques are being widely adopted by software community in various fields like bug finding and software verification.

References

1. Aho, A., Sethi, R., Ullman, J.: *Compilers: Principles, Techniques, Tools*. Addison Wesley, Boston (1986)
2. Nielson, F., Nielson, H.R., Hankin, C.: *Principles of Program Analysis*. 1st edn, p. 452, Springer, Berlin. (1996) ISBN: 978-3-540-65410-0
3. Kirkov, R., Agre, G.: *Source Code Analysis: An Overview*. Cybernetics and Information Technologies, Bulgarian Academy of Sciences (2010)
4. Ball, T.: The concept of dynamic analysis. In: Proceedings of 7th ACM/SIGSOFT Conference on Foundations of Software Engineering (1999)
5. Binkley, D.: *Source Code Analysis: A Road Map*. Future of Software Engineering, pp. 15–30. Minneapolis, USA, 23–25 May 2007
6. Ernst, M.D.: Static and dynamic analysis: synergy and duality. In: Proceedings of the 5th ACM Workshop on Program Analysis for Software Tools and Engineering (2004)
7. Gosain, A., Sharma, G.A.: Survey of dynamic program analysis techniques and tools. In: Proceedings of 3rd International Conference on Frontiers in Intelligent Computing Theory and Applications, Bhubaneshwar, vol. 1, pp. 113–122 Nov (2014)
8. Bentonino, A.: Software testing research: achievements, challenges, dreams. Future Softw. Eng. (2007)
9. Emaneulsson, P., Nilson, U.: A comparative study of industrial static analysis tools. Electron. Notes Theor. Comput. Sci. **217**, 5–21 (2008)
10. Brat, G., Venet, A.: Precise and scalable static program analysis of NASA flight software. In: IEEE Aerospace Conference, March (2005)
11. Jackson, D., Rinard, M.: Software analysis: a road map. IEEE Trans. Softw. Eng. (2000)
12. D'Silva, V., Kroenig, D., Weissenbacher, G.: A survey of automated techniques for formal software verification. IEEE Trans. CAD (2008)
13. Cifuentes, C.: BegBunch—benchmarking for C bug detection tools. DEFECTS (2009)
14. Pendergrass, J.A., Lee, S.C., McDonnell, C.D.: Theory and practice of mechanized software. Johns Hopkins APL Technical Digest, **32**(2) 2013
15. Kildall, G.A.: A unified approach to global program optimization. POPL (1973)
16. Kam, J.B., Ullman, J.D.: Global data flow analysis and iterative algorithms. J. ACM **23**(1), 158–171 (1976)
17. Kennedy, K.A.: Survey of data flow analysis techniques. In: Muchnick, S., Jones, N. (eds.) *Program Flow Analysis: Theory and Applications*, pp. 5–54. Prentice-Hall, Englewood Cliffs (1981)
18. Bush, W.R., Pincus, J.D., Sielaff, D.J.: A static analyzer for finding dynamic programming errors. Softw. Pract. Experience **30**(7), 775–802 (2000)
19. GrammaTech Inc. Overview of grammatech static analysis technology. White paper (2007)
20. Hovemeyer, D., Pugh, W.: Finding bugs is easy. <http://www.cs.umd.edu/~pugh/java/bugs/docs/findbugsPaper.pdf> (2003)
21. Das, M., Lerner, S., Siegel, M.: ESP: path sensitive program verification in polynomial time. PLDI'02, Berlin (2002)
22. Cousot, P., Cousot, R.: Abstract interpretation: a unified lattice model for static analysis of programs by construction or approximation of fixpoints. In Proceedings of 6th ACM Symposium on Principles of Programming Languages. California, pp. 238–252, (1977)
23. Jones, N.D., Nielson, F.: *Abstract Interpretation: A Semantics Based Tools for Program Analysis*. Handbook of Logics in Computer Science, vol. 14. Oxford University Press, Oxford (1995)
24. Cousot, P.: Abstract Interpretation Based Formal Methods and Future Challenges. Lecture Notes in Computer Science#2000, pp. 138–156. Springer, Berlin (2001)
25. Cousot, P., Cousot, R., Feret, J., Mine, A., Mauborgne, L., Monniaux, D., Rival, X.: Varieties of static analyzer: a comparison with astree. In: 1st Joint IEEE/IFIP Symposium on Theoretical Aspects of Software Engineering (TASE), June (2007)

26. Polyspace Verifier. <http://www.polyspace.com>
27. Aiken, A.: Introduction to set constraint-based program analysis. *Sci. Comput. Program.* **35**, 79–111 (1999)
28. Gulwani, S., Srivastava, S., Venkatraman, R.: Program analysis as constraint solving. *PLDI*, June (2008)
29. King, J.C.: Symbolic execution and program testing. *Commun. ACM* **19**(7), 385–394 (1976)
30. Floyd, R.: Assigning meanings to programs. In: *Proceedings of Symposium on Applied Mathematics* (1967)
31. Hoare, C.A.R.: An axiomatic basis for computer programming. *Commun. ACM* **12**(10), 576–580 (1969)
32. Detlefs, D., Nelson, G., Saxe, J.B.: Simplify: a theorem prover for program checking. www.hpl.hp.com/techreports/2003/HPL-2003-148.html
33. Tiwari, A., Gulwani, S.: Logical interpretation: static program analysis using theorem proving. In: *Proceedings of Conference on Automated Deduction* (2007)
34. Johnson, S.C.: Lint: A C program checker. Unix programmer's manual, Computer Science Technical Report 65. AT & T Bell Laboratories (1978)
35. FlexeLint/PCLint. <http://www.gimpel.com/html/lintinfo.htm>
36. Evans, D., Larochelle, D.: Improving security using extensible lightweight static analysis. *IEEE Softw.* **19**, 42–51 (2002)
37. JLint. <http://artho.com/jlint>
38. Flanagan, C., Leino, K.R.M., Lillibridge, M., Nelson, G., Saxe, J.B., Stata, R.: Extended static checking for Java. In: *Proceedings of ACM SIGPLAN Conference on Programming Language Design and Implementation*, pp. 234–245 (2000)
39. PMD/Java. <http://pmd.sourceforge.net>
40. Klocwork. Klocwork K7. <http://www.klocwork.com>
41. Chefl, B., Chou, A.: Next generation of static analysis: Boolean satisfiability and path simulation-a perfect match. http://www.coverity.com/library/pdf/coverity_white_paper_SAT_next_Generation_Static_Analysis.pdf. Downloaded on Dec 2012
42. Ami, T.L., Sagiv, M.: TVLA-a system for implementing static analyses. In: *Static Analysis Symposium* (2000)
43. Foster, J.S.: Type qualifiers: lightweight specifications to improve software quality. Ph.D. thesis, UCB (2002)
44. Aiken, A., Fahndrich, M., Foster, J., Su, Z.: A toolkit for constructing type- and constraint-based program analyses. In: *Proceedings of the 2nd International Workshop on Types in Compilation*, LNCS #, vol. 1473, pp. 76–96, March (1998)
45. Kodumal, J., Aiken, A.: Banshee: a scalable constraint- based analysis toolkit. In: *Proceedings of the 12th International Static Analysis Symposium*, pp. 218–234 (2005)
46. Ami, T.L., Reps, T., Sagiv, M., Wilhelm, R.: Putting static analysis to work for verification: a case study. *ISSTA* (2000)
47. Cousot, P., Cousot, R.: Compositional separate modular static analysis of programs using abstract interpretation. In: *Proceedings of 2nd International Conference on Advances in Infrastructure for E-Business, E-Science, E-Education on the Internet* (2001)
48. Dillig, T.: A modular and symbolic approach to static program analysis. Ph.D. Dissertation, Department of Computer Science, Stanford University (2011)

A LabVIEW-Based Data Acquisition System in a Quarter Car Test Rig to Optimize Vehicle Suspension System

Anirban C. Mitra and Nilotpal Banerjee

Abstract Suspension design is always been a challenging task for automobile designers in view of multiple input parameters, complex objectives, and disturbances which are stochastic in nature. The conflicting nature of ride comfort (RC) and road holding (RH) compels a judicial compromise between these two. In the present work, full factorial design of experiment (DoE) has been used successfully for the purpose of multi-objective optimization of RC and RH with input variables spring stiffness (K), damping coefficient (C), sprung mass (M), and speed. For experiments with combination of various input variables, a quarter car test rig is developed with proper data acquisition system by NI-hardware and Laboratory Virtual Instrumentation Engineering Workbench (LabVIEW) to collect real-time data. A regression model of RC and RH with R^2 value 95.98 and 95.78 %, respectively, can be effectively used to evaluate optimal settings of various input parameters using response optimization with a high desirability value.

Keywords Quarter car test rig · Ride comfort · Road holding · Design of experimentation · LabVIEW · DAQ system · Response optimization method

1 Introduction

The suspension system separates vehicle body frame from road surface. Harmful effects of vehicle vibrations on human body are nausea, back pain, osteoarthritis, slipping of disk, etc. [1]. The suspension systems are designed firstly to meet the ride comfort (RC) which is measured by root mean square (RMS) value of sprung

A.C. Mitra
M.E.S. College of Engineering, Pune, India
e-mail: amitra@mescoepune.org

N. Banerjee (✉)
National Institute of Technology, Durgapur, India
e-mail: nil_rec@yahoo.com

mass acceleration, a lower value which leads better RC and secondly, to satisfy, road holding (RH), which is measured by relative displacement between unsprung mass and road input, negative value of which leads to good RH [2]. It has been well established by other researchers in the field, that the ride characteristics of passenger vehicles can be characterized by considering a quarter car model of a vehicle [3]. This method is being widely used to investigate the performance of passive suspension system [4]. In this paper, optimization of RC and RH as per ISO 2631-1:1997 standard [5] is presented by the experimentation conducted on a quarter car test rig using the approach of design of experiments (DoE) response optimization method [6].

Laboratory Virtual Instrumentation Engineering Workbench (LabVIEW) is used to acquire data for RC and RH from quarter car test rig. IEPE accelerometers and proximity sensor are used to acquire readings for RC and RH, respectively [7, 8].

2 Data Acquisition Hardware and Sensors

In this study, multi-objective optimization of RC represents acceleration of sprung mass and RH represents relative displacement of road input and unsprung mass and rpm of CAM which represents the speed of vehicle approaching to road bump. Sensor selection was the most important step of a data acquisition system.

A high-sensitive 2-pin accelerometer of ICP® (IEPE) make (Model: 351B03) is used to measure the acceleration of sprung and unsprung mass.

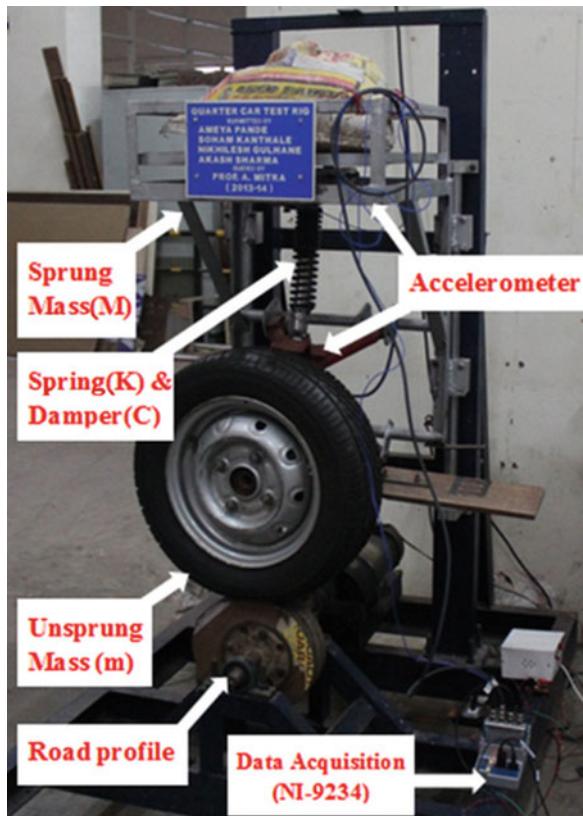
A proximity sensor (GP2Y0A21YK0F) [9] is used to measure the distance, which is an integrated combination of position-sensitive detector (PSD), infrared-emitting diode (IRED), and signal processing circuit. The variety of the reflectivity of the object, the environmental temperature, and the operating duration do not influence easily the distance detection because of adoption of the triangulation method. This device gives output as voltage corresponding to the detection distance.

To measure the rpm, an IR trans-receiver is placed facing each other so that extension connected to motor shaft may cut the rays between transmitter and receiver. Frequency as well as rpm of the shaft can be calculated by a LabVIEW programming. The entire sensors are connected to NI-9234 a 4-channel, ± 5 V, 24-Bit IEPE module. IEPE module is calibrated to all IEPE type of sensor, but proximity and IR sensor is non-IEPE-type sensor, which requires some scaling factors. Quarter car test rig with sensors and NI-9234 hardware is shown in Fig. 1.

3 LabVIEW Programming

LabVIEW is a platform and development environment for a visual programming language from National Instruments. LabVIEW coding is the most integral part of the data acquisition (DAQ) system. As LabVIEW is graphical coding language, it is

Fig. 1 Quarter car test rig with sensors and NI-hardware



easier to write code compared to other software. LabVIEW coding to acquire acceleration, displacement, and rpm is shown in Fig. 2, called ‘Block Diagram Panel.’ All the measured parameters can be displayed in a ‘front panel’ shown in Fig. 3. Time-domain acceleration data are also converted to frequency domain using a spectral measurement block.

NI-9234 hardware which is an integrated electronic piezoelectric (IEPE)-type module calibrated for ICP® IEPE-type accelerometers has been utilized for benefit. In this DAQ system, proximity sensor is also to be connected with NI-9234 by using external scaling factors mentioned in data sheet [9] as shown in Fig. 4.

Figure 5 represents a rpm measurement VI, where a frequency measurement block measures the frequency of interruption of the transmitting and receiving ray in the IR sensor which must be greater than 2 Hz (120 rpm) in real time, and for lower rpm measurement shows error. To tackle this problem, two extensions are attached on shaft so that measured rpm will be twice the actual rpm. A counter has been used to count the number of times the extension cuts the ray per second, and then, that value is converted in per minutes to measure rpm of the shaft.

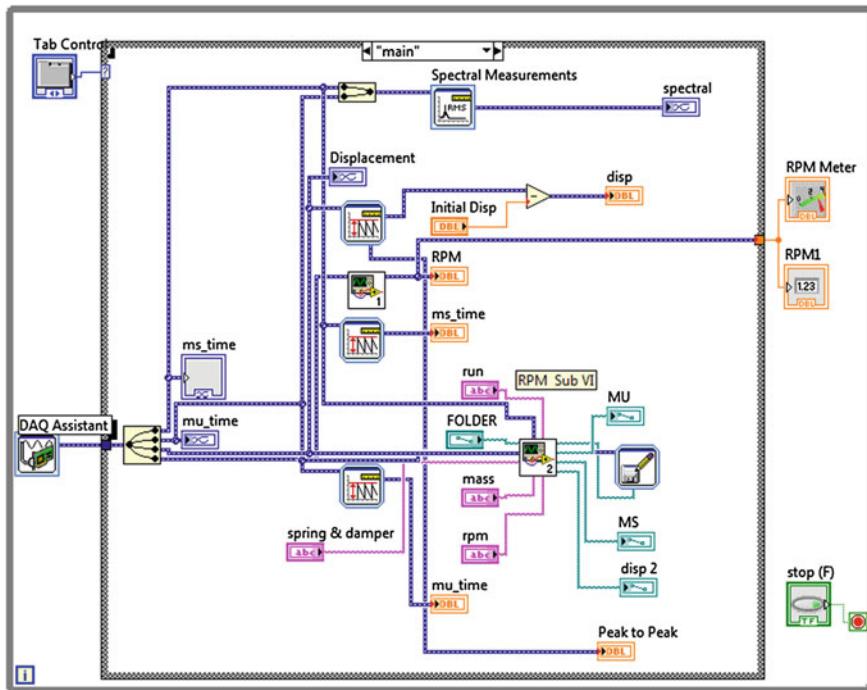


Fig. 2 LabVIEW coding for data acquisition of accelerometer, proximity, and IR sensor

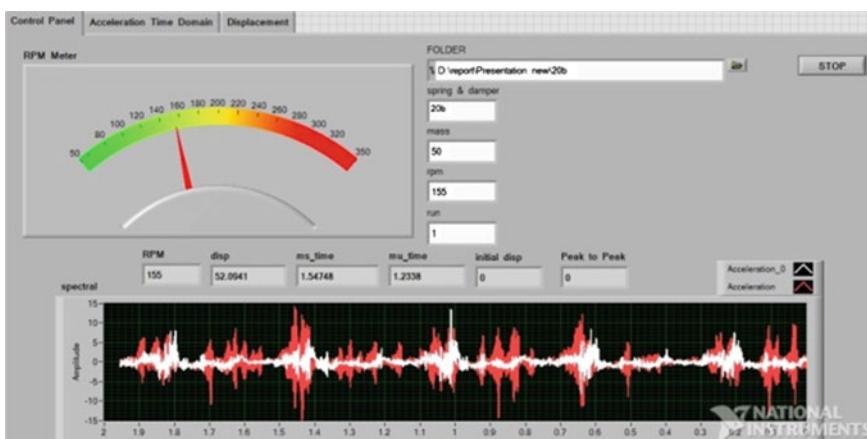


Fig. 3 Front panel of all sensors

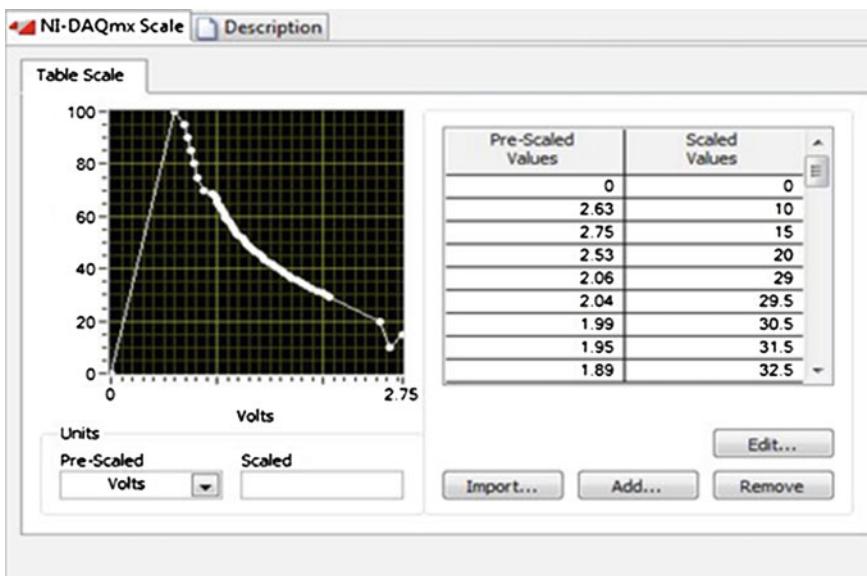


Fig. 4 Scaling factors of proximity sensor to measure displacement

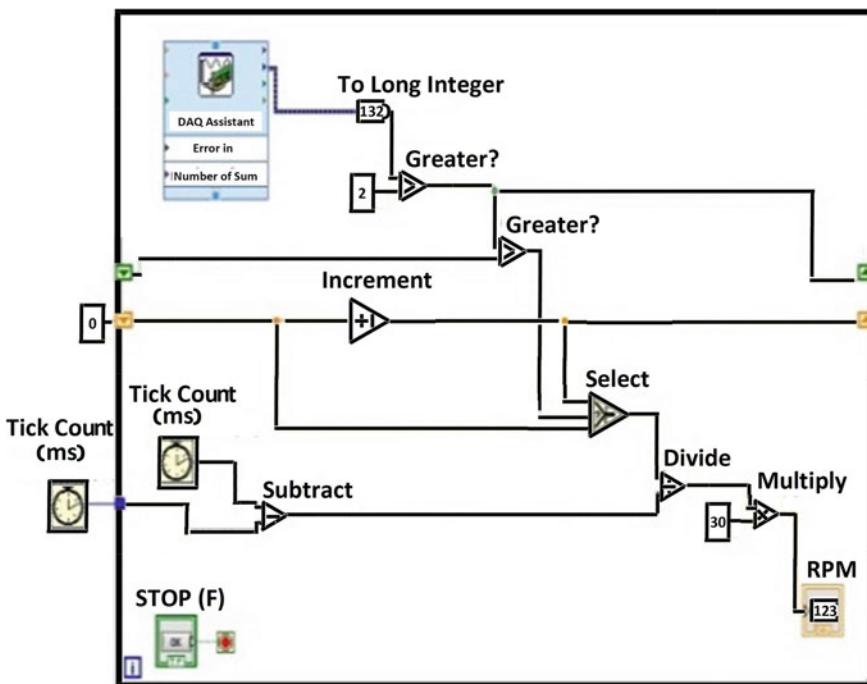


Fig. 5 LabVIEW VI for measuring rpm using IR sensor

4 Design of Experiments

The aim of the present work was to achieve multi-objective optimization formulations for RC and RH by determining the effects of some main factors and their interactions by regression analysis. DoE methodology was used for gathering information by conducting experiments on quarter car test rig (Table 1).

Four variables such as stiffness (K), damping coefficient (C), sprung mass (M), and speed are used as input variables with two-level and twice-replicating experimental reading, which are listed in Table 2 as per 2^k full factorial DoE matrix. The experiments are performed as per DoE matrix on the quarter car test rig, and values of various output parameters such as RC and RH are actually measured and stored in a computer using LabVIEW.

4.1 Regression Model

A MINITAB® software is used to fit the regression model with only significant terms. The ‘goodness of the fit’ of regression model in Eqs. (1) and (2) can be quantified from analysis of variance (ANOVA) by the R^2 value of RC 95.98 % and RH 95.78 %, which indicates the validity of model with input variables with 4.02 and 4.22 % variation due to error or by some noise factors, respectively. The high values of predicted R^2 as 89.70 and 89.21 % for RC and RH, respectively, ensures that models can be used for further analysis. The regression models with final significant estimated coefficients of RC and RH in uncoded units are as Eqs. (1) and (2).

$$\begin{aligned} \text{RC} = & 13.0909 + 0.000525529(K) + 0.0295497(C) + 0.123592(M) \\ & + 0.0124582(\text{Speed}) - 1.21068 \times 10^{-6}(K \times C) \\ & - 3.94993 \times 10^{-6}(K \times M) - 4.48506 \times 10^{-7}(K \times \text{Speed}) \\ & - 3.25464 \times 10^{-4}(C \times M) + 1.18596 \times 10^{-5}(C \times \text{Speed}) \\ & - 8.55855 \times 10^{-5}(M \times \text{Speed}) + 1.18198 \times 10^{-8}(K \times C \times M) \end{aligned} \quad (1)$$

Table 1 High and low level of input factors

Input factors	High level (+1)	Low level (-1)
Spring stiffness, K (N/m)	26,800	18,700
Damping coefficient, C (N-s/m)	673	418
Sprung mass, M (Kg)	90	50
Speed of CAM (rpm)	250	155

Table 2 Experimental readings as per full factorial DoE table

Run	<i>K</i> (N/m)	<i>C</i> (N-s/m)	<i>M</i> (kg)	Speed (rpm)	RC (m/s ²)	RH (cm)
1	18,700	418	90	250	0.63	-0.48
2	18,700	673	50	250	2.51	0.539
3	26,800	418	90	250	0.533	1.07
4	26,800	673	90	155	0.225	0.817
5	18,700	673	90	155	0.2727	-6.5
6	26,800	673	90	250	0.295	0.771
7	18,700	673	90	250	0.532	-4.27
8	18,700	673	50	155	1.4026	0.545
9	26,800	418	90	250	0.454	2.21
10	18,700	418	90	250	0.51	-0.98
11	26,800	418	50	155	0.5	-2.9
12	26,800	673	90	155	0.218	0.3505
13	26,800	418	50	155	0.731	-2.7
14	18,700	673	50	250	2.4	0.539
15	18,700	673	90	155	0.327	-6.66
16	18,700	418	50	250	1.1	1.15
17	18,700	673	50	155	1.42	0.552
18	26,800	673	50	155	0.339	-5.45
19	18,700	418	90	155	0.33	-0.8
20	26,800	673	90	250	0.296	0.28
21	26,800	673	50	250	0.596	-2.98
22	26,800	418	50	250	0.861	-1.2
23	26,800	673	50	155	0.354	-5.6
24	26,800	418	90	155	0.5	1.02
25	18,700	418	90	155	0.29	-0.6
26	26,800	418	50	250	0.762	-0.7
27	26,800	673	50	250	0.6167	-3.25
28	18,700	673	90	250	0.539	-4.4
29	18,700	418	50	155	0.79	0.48
30	18,700	418	50	250	0.87	0.966
31	26,800	418	90	155	0.85	2
32	18,700	418	50	155	0.801	0.82

$$\begin{aligned}
 \text{RH} = & -19.7075 + 0.000689108(K) + 0.101460(C) + 0.438444(M) \\
 & - 0.0051884(\text{Speed}) - 4.57799 \times 10^{-06}(K \times C) \\
 & - 1.41410 \times 10^{-05}(K \times M) + 5.5823 \times 10^{07}(K \times \text{Speed}) \\
 & - 0.00176016(C \times M) + 2.30057 \times 10^{05}(C \times \text{Speed}) \\
 & - 1.56036 \times 10^{-04}(M \times \text{Speed}) + 7.08893 \times 10^{-08}(K \times C \times M)
 \end{aligned} \tag{2}$$

5 Result and Discussion

Multiple response functions are estimated as polynomial functions of the factors in Eqs. (1) and (2) from regression analysis. The response variable is transformed to give a composite desirability value which is proportional to the priority given to the response variable. In this study, goals are set in Table 3 as per ISO 2631-1:1997 standard of RC with higher value of weightage and importance than RH.

The optimization plot in Fig. 6 shows the factor settings to get individual desirability [10] of both the RC and RH as 1.0. Therefore, the combined or composite desirability of these two variables is 1.0. To obtain this desirability, one would set the factor levels at values shown under global solution, i.e., $K = 18,700 \text{ N/m}$ and $C = 435 \text{ N-s/m}$, $M = 75 \text{ kg}$, and speed of CAM = 155 rpm.

Table 3 Set values of RC and RH for optimization

Parameter	Goal	Lower	Target	Upper	Weight	Importance
RC (m/s^2)	Target	0.315	0.5	1	0.6	1
RH (cm)	Target	-1.000	-0.5	0.00	0.4	2

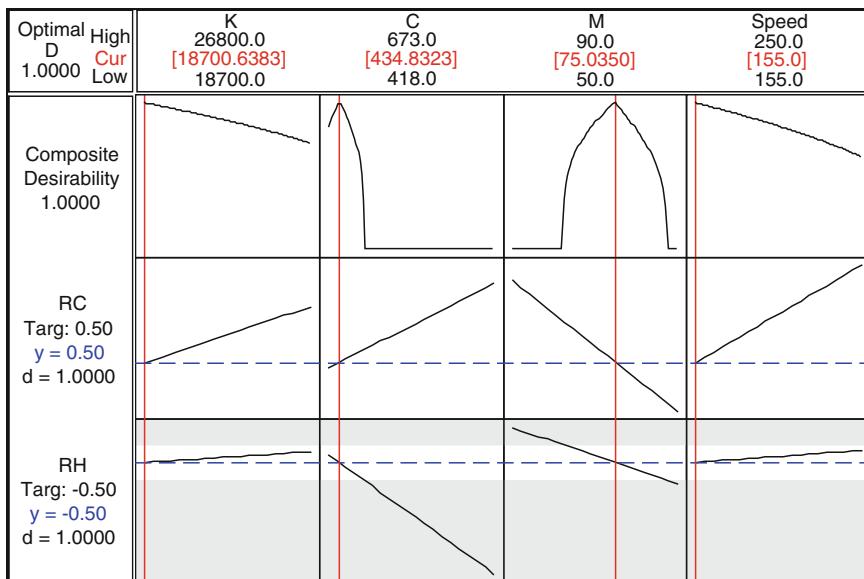


Fig. 6 Response optimization plot

6 Conclusion

In this study, a quarter car test rig is developed with proper DAQ system using LabVIEW. A regression model from this experimental results based on 2^k full factorial has been developed with significant terms of coefficient. The ‘goodness of fit’ for RC and RH models is checked from quantitative measures of R^2 as equal to 95.98 and 95.78 % with the R_{Pred}^2 equal to 89.70 and 89.21 %, respectively, being verified by the assumptions of random error by ANOVA. Such a model not only assists to estimate the magnitude and direction of the effects of change in the factors but also predicts the effects of their mutual interactions.

The model has assisted optimizing the RC and RH for factor setting of $K = 18,700 \text{ N/m}$, $C = 435 \text{ N s/m}$, $M = 75 \text{ kg}$, and speed = 155 rpm. With optimized parameters, the predicted values of RC and RH come out to be equal to 0.5 m/s^2 and -0.5 m , which is very much acceptable by ISO 2631-1:1997.

References

1. Kjellberg, A.: Psychological aspects of occupational vibration. *Scand. J. Work Environ. Health* **16**(Suppl. 1), 39–43 (1990)
2. Dixon, J.C.: *Suspension Geometry and Computation*. Wiley, New York (2009)
3. Crolla, D.A.: Vehicle dynamics: theory into practice. *Proc. Inst. Mech. Eng. Part D: J. Automobile Eng.* **210**, 83–94 (1996)
4. Sharp, R.S., Hassan, S.A.: Evaluation of passive automotive suspension systems with variable stiffness and damping parameters. *Veh. Syst. Dyn.* **15**(6), 335–350 (1986)
5. ISO: 2631-1, 1997: Mechanical vibration and shock—evaluation of human exposure to whole-body vibration (1997)
6. Montgomery, D.C.: *Design and Analysis of Experiments*, 7th edn. Wiley, New Jersey (2009)
7. Paton, B.E.: *Sensors, Transducers, and LabVIEW (Virtual Instrumentation Series)*, 1st edn. Prentice Hall PTR, New Jersey (1998)
8. Gupta, S., John, J.: *Virtual Instrumentation Using LabVIEW*. Tata McGraw Hill, New Delhi (2005)
9. Data sheet No: E4-A00201EN Date Dec.01.2006 ©SHARP Corporation (2006)
10. Designing Experiments Manual ©Minitab Inc. (2005)

A Computational Structural Biology of SoxR and DNA: A Modelling and Interactive Discern to Express the Sox Operon in *Pseudaminobacter salicylatoxidans* (KCT001) for Global Sulphur Oxidation

Sujay Ray, Arundhati Banerjee and Angshuman Bagchi

Abstract Computational and microbial molecular-level participation of sox operon and its repressor protein (SoxR) in sulphur oxidation from *Pseudaminobacter salicylatoxidans* (KCT001) was investigated. Documentation reveals that *P. salicylatoxidans* (KCT001) has sox *TRSVWXYZABCD* operon that is regulated by a repressor protein (SoxR). Previously, various experimental procedures such as DMS-mediated DNA methylations and hydroxyl radical footprinting have disclosed that SoxR interacts first with an operator region-*sv* (present in between soxS and soxV). Detailed computational studies were accomplished in the present study. 3D models of repressor protein and the DNA sequence from operon's promoter region were demonstrated using molecular modelling techniques. Molecular docking simulation was performed to predict DNA–protein interaction. Amino acid residues and nucleotide bases responsible for interaction were identified by PyMOL and Discovery Studio software suite. This novel residue-level study is paramount for initiating transcription in the operon, thereby leading to sulphur oxidation.

Keywords Sulphur oxidation · Thiosulphate · Sox operon · SoxR · Homology modelling · Docking simulations

S. Ray · A. Bagchi (✉)
Department of Biochemistry and Biophysics, University of Kalyani,
Kalyani, Nadia, India
e-mail: angshuman_bagchi@yahoo.com; angshumanb@gmail.com

S. Ray
e-mail: raysujay@gmail.com

A. Banerjee
Department of Biotechnology, National Institute of Technology,
Mahatma Gandhi Avenue, Durgapur, West Bengal, India
e-mail: arundhati.92star@gmail.com

Abbreviations

- SD Steepest Descent
CG Conjugate Gradient

1 Introduction

It is a prime necessity to oxidize the sulphide to elemental sulphur as hydrogen sulphide is highly toxic to plants and animals. Thiosulphate is universally oxidized by majority of known chemolithotrophic organisms. It oxidizes directly to sulphate by α -proteobacteria. Earlier genetic studies with both chemo- and photolithotrophic sulphur-oxidizing α -proteobacteria have disclosed a cluster of sulphur oxidation (sox) genes, viz. *soxVW* and *soxXYZABCDEFGH* [1, 2]. α -Proteobacteria: a model system having a sulphur-oxidizing multienzyme complex comprises thiosulphate-induced periplasmic proteins SoxXA, SoxYZ, SoxB, and SoxCD [3–5]. Documentation showed SoxV and SoxW being involved in biosynthesis or maintenance of multienzyme complex system [2, 6]. This α -proteobacteria model system follows a consensus mechanism for the complete global sulphur oxidation. SoxR, an ArsR/SmtB family protein [7] is encoded by ORF *soxR*, present in differently transcribed *soxSRT* gene cluster [8].

Repression of *sox* gene by SoxR in *Pseudaminobacter salicylatoxidans* (KCT001) and *Paracoccus pantotrophus* is mediated by the repressor-operator interaction [6]. Structurally, SoxR protein belongs to winged helix family, having a helix-turn-helix (HTH) domain [9]. SoxR repressor protein regulates chemolithotrophic *sox* operon by encoding a multienzyme complex responsible for complete oxidation of thiosulphate to sulphate from *soxSRT-soxVWXYZABCD* gene cluster [10, 11]. Earlier wet-laboratory documentations through experimental procedures such as dimethyl sulphate (DMS)-mediated DNA methylations and hydroxyl radical footprinting [12] reveal that SoxR first binds to an operator in *sox* operon which is in between *soxS* and *soxV* (*sv*) genes. This *sv* site has two joint long adjacent symmetrical sites for their binding [9]. Certain experimental deductions [9] account that interaction with *sv* site further acts as an opening for SoxR to interact further with the next operator which is in between *soxW* and *soxX* (*wx*) genes. This *wx* site has two adjacent sites: a high-affinity one and a low-affinity one, separated by a pair of nucleotides. Among them, the high-affinity zone encounters the interaction first followed by the interaction of SoxR with the low-affinity zone of *wx*. This now leads to the entry into the *sox* operon for its expression leading to removal to sulphides from the environment. It is well documented that [13] at only low concentrations, SoxR is an efficient repressor.

The present study involves a residue and molecular-level study on *sox* operon, which thereby causes participation of gene encoding enzymes for oxidation of sulphur compounds. Current study includes demonstrative analysis of 3D structure of SoxR protein and the DNA structure in the *sv* site of the operator region by

molecular modelling techniques. Investigation in detailed SoxR-*sv* site interaction pattern and conformational changes, if any, is essential to further study the sequential and consecutive interaction by SoxR for entering the sox operon to create a sulphide-free and toxic-free surrounding. Analysis of individual binding patterns and molecular-level interactions was performed by molecular docking simulations for this initial major interaction. Discovery Studio platform from Accelrys and PyMOL showed the involvement of amino acid residues and nucleotide sequences.

This being most probably an unexplored and novel study; it provides clear information about molecular-level interactions between SoxR protein and nucleotide sequences, in the presence of thiosulphate. It is essential for initiating transcription, thereby leading to sulphur oxidation.

2 Materials and Methods

2.1 Sequence Analysis and Template Search

The amino acid sequence of SoxR protein of *P. salicylatoxidans* (KCT001) from GenBank has Accession No.: AJ404005.4. Results from BLASTP against PDB and HHpred inferred that SoxR had template in X-ray crystal structure from *Vibrio cholereae* (PDB code: 4K2E_A) sharing 42 % sequence identity. HHpred is highly sensitive method, which deduces inferences from more remotely homologous relationships [14].

2.2 Molecular Modelling of SoxR Protein of *P. salicylatoxidans* KCT001

After obtaining the sequence of SoxR, homology modelling with MODELLER was performed for SoxR using the sequence as it satisfies the prerequisite to have sequence identity of more than 30 % with the templates [15]. To study the globular protein conformation, RMSD value of 0.843 Å was obtained by superimposing Cα atoms of modelled target structure on its respective original template.

2.3 Loop Optimization Using ModLoop

ModLoop optimized and remodelled SoxR protein structure for $\psi-\varphi$ angle conformation as certain loop regions occupied disallowed regions in Ramachandran Plot [16]. It performs automated loop modelling of protein structures using MODELLER.

2.4 Molecular Modelling of the DNA Sequence from sv Site

SoxR protein first binds to the operator in a DNA sequence having nucleotide sequence: GTTGTGTATAGTAATAGACGTATGTATGCGG in 3'-5' direction for sv site. It was modelled using Discovery Studio 2.5 software packages. B-DNA form was utilized further.

2.5 Energy Minimization of SoxR and DNA Model

Energy minimization of SoxR and the DNA was performed using CHARMM force field. Thus, the overall energy was minimized [17]. In minimization phenomena, steepest descent technique was followed by conjugate gradient using Discovery Studio software, until they reached final RMS gradient of 0.0001 [17].

2.6 Validation of Model

VERIFY3D was used to check amino acid residue profile of SoxR model [18]. Main chain properties of modelled SoxR were verified by SAVeS server. PRO-CHECK Web server showed that there were no amino acid residues in the disallowed region of Ramachandran Plot [19, 20].

2.7 Molecular Docking Studies of DNA with SoxR

Binding of SoxR with the DNA from sv site was executed using PatchDock server where output comprises a list of potential complexes sorted by shape complementarity criteria. PatchDock: An algorithm for molecular docking [21] accepts two molecules of any type: proteins, DNA, peptides, and drugs. Surfaces of the two molecules are divided into patches according to the surface shape which are superimposed using shape matching algorithms. The molecular docking results were also validated from Hex Server [22].

2.8 Energy Minimization of DNA–Protein Docked Complex

The docked DNA–protein complexes were also energy-minimized using CHARMM force field. Energy minimization for docked complex structures was performed firstly by allowing flexibility in all main and side chain atoms of protein

and then side chains were allowed to remain flexible for proper interactions keeping backbones of protein fixed. Energy minimization was done by steepest descent technique first, followed by the conjugate gradient using Discovery Studio software by CHARMM force field until they reached the final RMS gradient of 0.0001 [17].

3 Results

3.1 Structural Description of SoxR

Modelled structure of SoxR protein from *P. salicylatoxidans* KCT001 is 121 amino acid residues long (Fig. 1). Predicted structure was found similar to the structure of *V. cholereae* (PDB code: 4K2E_A). It comprises six helical structures (residues: 18–22, 27–39, 41–52, 57–61, 69–81, and 100–113) interspersed with two anti-parallel beta-sheets (residues: 84–89 and 92–97) and coil regions.

3.2 Structural Description of sv Site

DNA structure from sv site was found to have nucleotide sequence: (3'GTT GTGTATAGTAATAGACGTATGTATGCGG5'), in cyan shade, coiled with their complementary strands (magenta shade) 5'CAACACATACATTATCTGCATA CATA CGCC3' (Fig. 2). Major and minor grooves were clearly visible.

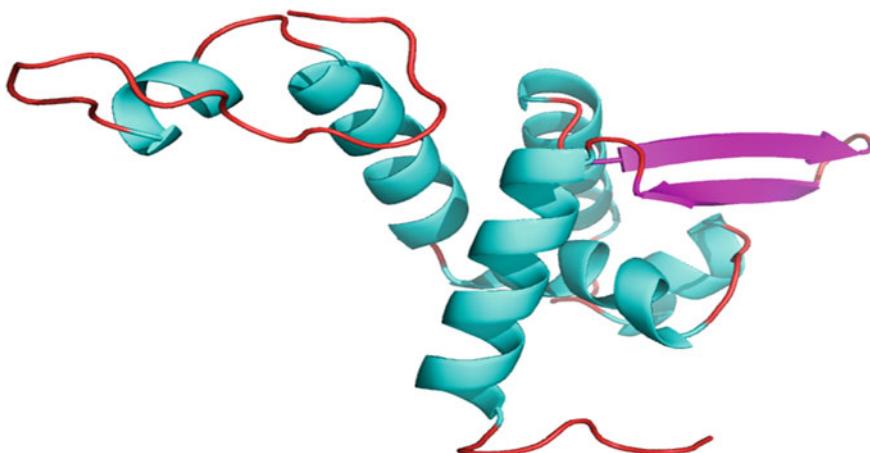


Fig. 1 Modelled SoxR showing distinct helices (cyan), sheets (magenta), and coils (red)



Fig. 2 Modelled DNA from *sv* site having prominent major and minor grooves with complimentary strands in magenta shade

3.3 Interaction of SoxR with *sv* Site

The 3D coordinates of the SoxR–DNA complex has been generated using molecular docking technique. Previous experimental phenomena such as DMS-mediated DNA methylations and hydroxyl radical footprinting [13] reveal that SoxR binds first to an operator in *sv* site. Documentation [9] accounts that *sv* site has two joint adjacent symmetrical sites without gaps.

SoxR and the DNA are bound together by strong non-covalent forces. The predominant interactions are the hydrogen-binding (H-bonding) interactions. Residues from SoxR such as Leu9, Ile97, Asn85, Glu101, Lys120, and Arg118 form strong H bonds with C25 (antisense strand), T18 (antisense strand), T17 (antisense strand), T18 (antisense strand), G1 (sense strand), C20 (antisense strand), and A19 (antisense strand) of the DNA (*sv* site), respectively. The two thiamine residues in 17th and 18th positions of DNA antisense strand formed three H bonds with Leu9, Ile97, and Asn85. Thus, a stronger non-covalent interaction is formed. In the protein, a cavity that is most probably formed due to the charged polar residues such as Glu101, Lys120, and Arg118 facilitates the binding of DNA to it. Furthermore, Lys120 and Arg118 formed a strong positive environment around the DNA-binding region. This aids in the initiation of transcription in sox operon and global sulphur oxidation system. The interactions are shown clearly in Fig. 3 and Table 1.

4 Discussion

The aim of the study was to analyse the probable biochemical role and residue-level participation of SoxR repressor protein from sox operon of *P. salicylatoxidans* (KCT001), in global sulphur oxidation reaction cycle. It has been discerned from the previous experimental studies that only in lower concentrations, SoxR acts as an efficient repressor by binding to an operator: *sv* site. From ArsR/SmtB family protein [7], SoxR is a DNA-binding protein comprising helix–turn–helix domain. In order to predict the mode of binding of SoxR with the DNA in the *sv* site of the operator, 3D structures of SoxR protein and DNA have been generated using

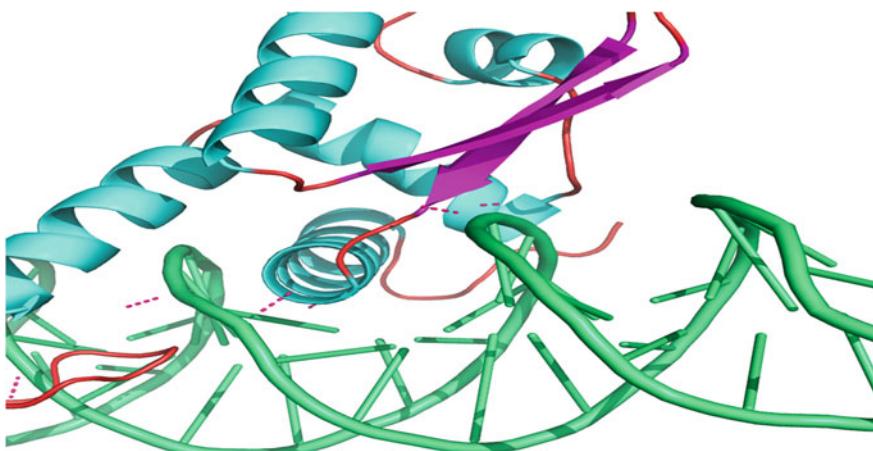


Fig. 3 Interaction through H bonds (shown in *magenta dashes*) between SoxR and sv-site DNA

Table 1 The interacting amino acid residues and nucleotide bases along with their individual positions at the binding interface

DNA–protein complex	DNA residue number	Protein residue number	H bonds
SoxR-sv site complex	Antisense (a) strand: ATTACTACTATACA (16–26)	4–9,11,18–19, 22, 27, 30, 31, 34–36, 38, 49–54, 85, 87–88, 95–109, 111, 112, 118, 120	LEU9:aC25 ASN85:aT17 ILE97:aT18 GLU101:sG11 LYS120:aC20 ARG118:aA19
	Sense (s) strand: TATAGTA (7–13), GAC (17–19)		

comparative modelling technique. Molecular docking simulation studies were then conducted to observe the interaction between protein–nucleotide complexes along with the binding patterns. Amino acid residues and nucleotide bases involved in the phenomena were identified through Discovery Studio software packages and PyMOL. The putative mode of interaction and binding patterns from the current study revealed that the predominant interactions were the H-bonding interactions (Table 1 and Fig. 3).

In the protein, charged polar residues such as Glu101, Lys120, and Arg118 form a cavity that facilitates the binding of DNA to it. Among which, Lys120 and Arg118 formed a strong positive environment around the DNA-binding region. Interestingly, along with rest of the residues involved as shown in Table 1 and Fig. 3, the two thymine residues in 17th and 18th positions of the antisense strand aided in an additional stability to the SoxR-sv site complex structure.

Previously, no reports were documented, involving computational study of repressor protein and the sv-site DNA from sox operon of *P. salicylatoxidans*

(KCT001) for sulphur oxidation. So, results from the current study provide a rational framework for designing experiments to determine indulged residues and nucleotide bases, essential for initiation of transcription in sox operon and thus sequentially leading to sulphur oxidation and thereby a balanced environment.

5 Conclusion

SoxR-sv site interaction in the sox operon is one of the paramount zones that needed to be investigated, as it holds the preliminary aspect for getting rid of the sulphide from being in the toxic environment. It is only after this specific interaction the SoxR gets an opportunity to make a vivid entry into the sox operon via the interaction with *wx* site of the operator. This leads to the expression of several structural genes in the sox operon that aids in global microbial sulphur oxidation. It leads to complete removal of sulphides, making a sulphide-free environment.

As, previously, there was no such examination that dealt with the gateway for sulphur oxidation, so this study serves a prime requisite and might be brought to limelight for the purpose. In future, interaction of SoxR with *wx* site and its outcomes can be discerned. This study may also serve as a template for studying the molecular level of interaction to remove few more pollutants from the environment making it a better place to live in.

Acknowledgment Authors are deeply indebted to DST-PURSE program 2012–2015 and DBT sponsored Bioinformatics Infrastructure Facility in Department of Biochemistry and Biophysics, University of Kalyani, for providing different equipments and essential infrastructural support.

Conflicts of interest None

References

1. Friedrich, C.G., Quentmeier, A., Bardischewsky, F., Rother, D., Kraft, R., Kostka, S., Prinz, H.: Novel genes coding for lithotrophic sulfur oxidation of *Paracoccus pantotrophus* GB17. *J. Bacteriol.* **182**, 4677–4687 (2000)
2. Appia-Ayme, C., Berks, B.C.: SoxV, an orthologue of the CcdA disulfide transporter, is involved in thiosulfate oxidation in *Rhodovulum sulfidophilum* and reduces the periplasmic thioredoxin SoxW. *Biochem. Biophys. Res. Commun.* **296**, 737–741 (2002)
3. Friedrich, C.G.: Physiology and genetics of sulfur-oxidizing bacteria. *Adv. Microb. Physiol.* **39**, 235–289 (1998)
4. Friedrich, C.G., Rother, D., Bardischewsky, F., Quentmeier, A., Fischer, J.: Oxidation of reduced inorganic sulfur compounds by bacteria: emergence of a common mechanism? *Appl. Environ. Microbiol.* **67**, 2873–2882 (2001)
5. Appia-Ayme, C., Little, P.J., Matsumoto, Y., Leech, A.P., Berks, B.C.: Cytochrome complex essential for photosynthetic oxidation of both thiosulfate and sulfide in *Rhodovulum sulfidophilum*. *J. Bacteriol.* **183**, 6107–6118 (2001)

6. Bardishevsky, F., Friedrich, C.G.: The shxVW locus is essential for oxidation of inorganic sulfur and molecular hydrogen by *Paracoccus pantotrophus* GB17: a novel function for lithotrophy. *FEMS Microbiol. Lett.* **202**, 215–220 (2002)
7. Busenlehner, L.S., Pennella, M.A., Giedroc, D.P.: The SmtB/ArsR family of metalloregulatory transcriptional repressors: structural insights into prokaryotic metal resistance. *FEMS Microbiol. Rev.* **27**, 131–143 (2003)
8. Mandal, S., Chatterjee, S., Dam, B., Roy, P., Das Gupta, S.K.: The dimeric repressor SoxR binds cooperatively to the promoter(s) regulating expression of the sulfur oxidation (sox) operon of *Pseudaminobacter salicylatoxidans* KCT001. *Microbiology* **153**, 80–91 (2007)
9. Bagchi, A., Roy, D., Roy, P.: Homology modeling of a transcriptional regulator SoxR of the lithotrophic sulfur oxidation (Sox) operon in α -proteobacteria. *J. Biomol. Str. Dyn.* **22**, 571–578 (2005)
10. Mukhopadhyaya, P.N., Deb, C., Lahiri, C., Roy, P.: A soxA gene, encoding a diheme cytochrome c, and a sox locus, essential for sulfur oxidation in a new sulfur lithotrophic bacterium. *J. Bacteriol.* **182**, 4278–4287 (2000)
11. Mandal, S., Ghosh, W., Dam, B., Roy, P.: A novel gene cluster soxSRT is essential for the chemolithotrophic oxidation of thiosulfate and tetrathionate by *Pseudaminobacter salicylatoxidans* KCT001. *Curr. Microbiol.* **54**, 267–273 (2006)
12. Mandal, S., Das Gupta, S.K.: Interactions of sulfur oxidation repressor with its promoters involve different binding geometries. *Arch. Microbiol.* **194**, 737–747 (2012)
13. Mandal, S., Chatterjee, S., Dam, B., Roy, P., Das Gupta, S.K.: The dimeric repressor SoxR binds cooperatively to the promoter(s) regulating expression of the sulfur oxidation (sox) operon of *Pseudaminobacter salicylatoxidans* KCT001. *Microbiology* **153**, 80–91 (2007)
14. Söding, J., Biegert, A., Lupas, A.N.: The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33**, W244–W248 (2005)
15. Sander, C., Schneider, R.: Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins* **9**(1), 56–68 (1991)
16. Fiser, A., Sali, A.: ModLoop: automated modeling of loops in protein structures. *Bioinformatics* **19**(18), 2500–2501 (2003)
17. Brooks, B.R., Brucoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., Karplus, M.: CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* **4**(2), 187–217 (1983)
18. Eisenberg, D., Luthy, R., Bowie, J.U.: VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods Enzymol.* **277**, 396–404 (1997)
19. Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M.: PROCHECK: a program to check the stereochemistry of protein structures. *J. Appl. Cryst.* **26**, 283–291 (1993)
20. Ramachandran, G.N., Sasisekharan, V.: Conformation of polypeptides and proteins. *Adv. Protein Chem.* **23**, 283–438 (1968)
21. Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., Wolfson, H.J.: PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucl. Acids. Res.* **33**, W363–W367 (2005)
22. Macindoe, G., Mavridis, L., Venkatraman, V., Devignes, M.-D., Ritchie, D.M.: HexServer: an FFT-based protein docking server powered by graphics processors. *Nucleic Acids Res.* **38**, W445–W449 (2010)

Empirical Evaluations Using Character and Word N-Grams on Authorship Attribution for Telugu Text

**S. Nagaprasad, T. Raghunadha Reddy, P. Vijayopal Reddy,
A. Vinaya Babu and B. VishnuVardhan**

Abstract Authorship attribution (AA) is the task of identifying authors of anonymous texts. It is represented as multi-class text classification task. It is concerned with writing style rather than topic matter. The scalability issue in traditional AA studies concerns with the effect of data size, the amount of data per candidate author. Most stylometry researches tend to focus on long texts per author, but it is not probed in much depth in short texts. This paper investigates the task of AA on Telugu texts written by 12 different authors. Several experiments were conducted on these texts by extracting various lexical and character features of the writing style of each author, using word n-grams and character n-grams as a text representation. The support vector machine (SVM) classifier is employed in order to classify the texts to their authors. AA performance in terms of F_1 measure and accuracy deteriorates as the number of candidate author's increases and size of training data decreases.

Keywords Authorship attribution · Telugu language · Support vector machine · Evaluation measures · Word n-grams · Character n-grams · Text classification

S. Nagaprasad (✉)

Department of CSE, Acharya Nagarjuna University, Guntur, India
e-mail: nagkanna80@gmail.com

T. Raghunadha Reddy

Department of CSE, Swarnandhra Institute of Engineering and Technology, Narsapur, India

P. Vijayopal Reddy

Department of CSE, Gokaraju Rangaraju Institute of Engineering and Technology,
Hyderabad, India

A. Vinaya Babu

Department of CSE, J.N.T.U. College of Engineering, Hyderabad, India

B. VishnuVardhan

Department of IT, J.N.T.U. College of Engineering, Nachupally, Karimnagar, India

1 Introduction

Authorship attribution (AA) is the technique of determining the author of a text when it is ambiguous to identify the real author [1]. Every author has his own writing style. Invariably, AA applications are plagiarism detection, resolving disputed authorship, criminal law, civil law and data security [2]. AA can be viewed as problem of text classification (TC), but it is different from classification in terms of considering embedded author style in addition to text content. Hence, AA is more challenging compared with text classification. In TC, the problem is viewed as identification of related topic or the given test document, whereas in AA, it is viewed as assigning the test document to one the many predefined authors [1]. AA is a research field that is in the last decade on various data sets of various languages. AA is identified as a stylometry problem [3] till computational techniques were enough matured. Usage of computational techniques in AA gives pathway for considering many other aspects other than linguistic features. Different data sets with various sizes were experimented in combination with different features and with different machine learning algorithms. Based on the data set size and the based on the number of authors, features and machine learning approaches behave differently [4]. In order to evaluate the proposed AA method thoroughly, its performance is measured under various conditions [5] such as training corpus size and number of candidate authors. Unless these issues are addressed, it is impossible to claim superiority of any type of features for AA. In this paper, a systematic study of the features of AA, such as effect of author set size; data size on performance; and influence of lexical and character features in a categorization approach using support vector machine (SVM) was presented. The behaviour of SVM and the predictive strength of different types of features using various author sets sizes and varying data sizes on Telugu data set were compared. To our knowledge, this is the first study of these aspects of AA on Telugu data set.

2 Related Work

AA can be viewed as one of the oldest problems and one of the newest research problems in the field of information retrieval. Stylometry is the statistical analysis of literary style. The main assumption behind stylometry is that the authors make certain subconscious and conscious choices in their writing. Some of the features that were used in stylometry include average sentence length, average syllables per word, average word length, distribution of parts of speech, function word usage, the type–token ratio, Simpson’s Index, Yule’s Characteristic K, entropy, word frequencies and vocabulary distributions [6]. Some models that were used in stylometry include n-grams [7], feature counts, inductive rule learning, Bayesian networks, radial basis function networks, decision trees, nearest neighbour classification and SVM [8]. Mosteller and Wallace [9] propose to select semiautomatically the most

common terms composed mainly by various function words for AA. The earliest studies of AA were reported by Yang et al. [10] and Yule [11], in which statistical methods were used limit data, not only the size of the experimental corpus but also the size of feature set. Yang et al. [10] graphically represented the word length as characteristic curves, and he also in [11] used sentence length to differentiate between authors text.

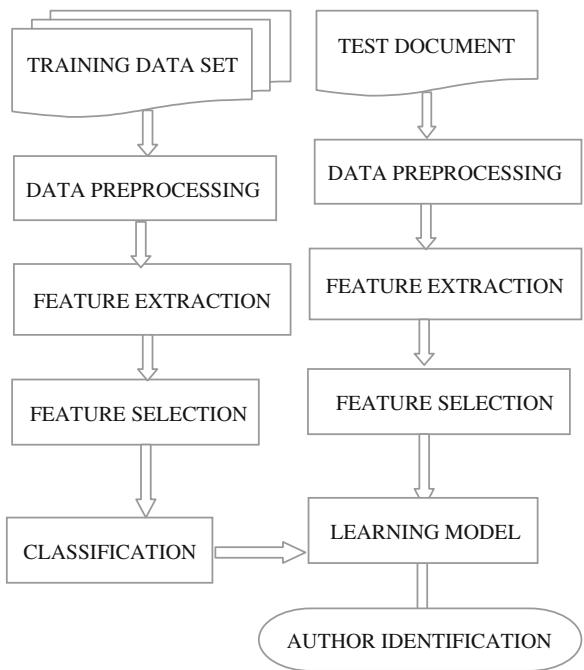
Many types of lexical features were proposed [12–16] including token-level style markers, frequency of word usage, richness of the vocabulary, including the distribution of vocabulary, the number of hapax legomena. Burrows [17] extracted commonly used words from the collection as the features. Stamatatos in [18] was experimented with vocabulary richness. Stamatatos also pointed out in [19] that merely using features at the token level are not be sufficient for reliable AA. Burrows in [20] first proposed the use of function words as style markers for AA. Baayen in [13] experimented with 42 common function words and eight punctuation symbols. A set of 50 common function words was selected as style markers by Holmes in [15] in order to discriminate between two authors on disputed journal articles. Binongo in [21] also used 50 common function words to examine the authorship. More function words were used by Juola and Baayen [16]. Pol [22] has carried out experiments to discriminate the power of different lexical features. Grammatical-based or syntax-based features in AA were applied by several researchers [13, 23]. Chi-square (χ^2) measure is often used to determine relevant features in authorship attribution [19, 24]. The cumulative sum technique [25] looks at the frequencies of a range of possible habits in use of language. Principal component analysis (PCA) [15–17], Markov chains [13] and compression-based techniques [26] are typical of computational approaches that were proposed for AA. N-grams are widely used in authorship attribution [18, 27]. Juola in [28] proposed a similar approach that was applied to AA, in which the unigram model on the character level was used. Benedetto in [26] used compression approach to different applications including AA. Machine learning approaches were applied to AA in recent years, including neural networks [29], Bayesian classifiers [30], SVMs [31] and decision trees [32].

In this paper, Sect. 3 discusses about the different steps in proposed model such as data preprocessing, feature extraction, feature selection and machine learning approach. The AA accuracy and F_1 measure in terms of data set size and author set size were evaluated in the Sect. 4. The description about the data set collection and the language characteristics was presented in Sect. 4. Section 5 summarizes the work done, and the conclusions were drawn from the results and possible extensions.

3 Author Attribution Model

Authorship attribution is viewed as an automatic text classification task that assigns documents according to a set of predefined author set. AA model consists of various steps as shown in Fig. 1. They are data preprocessing, feature extraction and feature selection. These three steps are performed on both training set and testing set.

Fig. 1 Text authorship attribution model



A learning model was generated for each author using machine learning technique SVM, and finally, test document is assigned to one of the known authors using the learning model.

3.1 Data Preprocessing

The raw text documents are not suitable for processing by the machine learning algorithms. In this scenario, there is a need to convert these raw documents into a suitable format such as attribute-value representation. This step contains the tokenization, stemming and stop word removal as in [33].

Tokenization is a process of dividing the raw text into meaningful elements. The elements are in the form of paragraphs, sentences, phrases, words and also characters. Based on the characteristics of the language, various types of elements may have various meanings. Tokens such as punctuation symbols, whitespace and numbers are not included in the input token list as they are not deriving any meaning to be extracted from the text. This token acts as an input data for the remaining steps in the proposed model.

To reduce the feature space of the token set, all the words are reduced to their stem form. Stemming is the process of deriving root or base form of the original word which is from the token set. The stemmed word may not be the root of the original word, but

all the related words are mapped to a single word known as stemmed word. For the Telugu data set used in the paper, the stemmed words are derived using the Telugu stemmer called Telugu morphological analyser (TMA) as in [34].

3.2 Feature Extraction

The style of a particular author is generally identified by extracting various features from the text. The various features of a text are broadly categorized into three types as in [32], namely lexical, syntactic and structural features. In this paper, lexical n-grams such as syllable n-grams and word n-grams as features were considered. Vishnu in [35] claimed that lexical n-grams are best features for Telugu text classification. As in [36, 37], lexical features are good for small data sets and also able to capture nuances in different linguistic levels: it is considered that the syllable unigram, syllable bigram, syllable trigram and syllable tetragram at syllable level and also word unigram, word bigram, word trigram and word tetragram as features at word level.

3.3 Feature Selection

The extracted features from the previous step may increase the dimensionality space of the input set. The machine learning classifiers suffer with the problem of curse of dimensionality as the dimensionality space increases. Hence, it is required to remove irrelevant or not most relevant features from the features set. It is carried out by various measures such as document frequency, DIA association factor, chi-square, information gain, mutual information, odds ratio, relevancy score and GSS coefficient. In this paper, chi-square (χ^2) metric [25] is used as a measure for feature selection, which is the most effective feature selection metric in the literature [26]. Chi-square measures the correlation between feature and author set. The relevance of feature t with the author set c is calculated as follows:

$$\chi^2(t, c) = \frac{N * (AD - BC)^2}{(A + C) * (B + D) * (A + B) * (C + D)}$$

where A is the number of times both feature t and author set c exist; B is the number of times feature t exists, but author set c does not exist; C is the number of times feature t does not exist, but author set c exists; D is the number of times both feature t and author set c does not exist; N be the total number of the training samples. As the value is more, the feature t is more relevant to the set c . Some of the features whose chi-square value is less than the threshold value are considered as non-relevant to the class c .

3.4 Classification

SVM is proved to be an effective machine learning algorithms for text categorization. In [38] for AA, SVM is used to generate learning model by using lexical features such as character n-grams and word n-grams to represent the text. SVM classifier is used to learn the boundaries between author sets where author sets are treated as classes. The learned model generated from the SVM is used for author identification of unknown text as shown in Fig. 1.

3.5 Author Identification

In this step, author is assigned for a given unknown text document. Unknown author text is processed through the various steps as shown in Fig. 1. The vector representation of the text document after dimensionality reduction is given as input to learn the model which is generated from the classifier. The learned model assigns one of the authors from author set to the test document.

4 Results and Discussions

The following Sect. 4.1 briefly describes the characteristics of the language; Sect. 4.2 describes the data set. The different evaluation measures used in authorship attribution are explained in Sect. 4.3; Sect. 4.4 discusses about the influence of the number of authors on AA. Sect. 4.5 presents the influence of the size of the data for each author on AA.

4.1 About the Language

There are more than 150 different languages spoken in India today. Indian languages are characterized by a rich system of inflectional morphology and a productive system of derivation. This means that the number of surface words found to be very large and so the raw feature space, leading to data scarcity. Dravidian languages such as Telugu and Kannada are morphologically more complex and comparable to the languages such as Finnish and Turkish as in [34]. The main reason for richness in morphology of Telugu (and other Dravidian languages) is the significant part of grammar that is to be handled by syntax in English (and other similar languages) to be handled within morphology. Phrases including several words in English are mapped on to a single word in Telugu. Hence, there is a necessity to study the influence of features and different AA approaches on Indian context.

4.2 Data set Description

To address the problem of authorship attribution on Indian context for Telugu language, the data set is collected from Telugu newspapers. The collected data cover various topics. The data set contains 300 news articles written by 12 authors. The average numbers of words are 547 per document. In our experiments, the data set is separated into two groups such as training and testing data. The training set contains 20 text articles for each author. On the other hand, for the test set consists of 5 text articles for each author. The training data set is used to generate learning model using SVM algorithm. Each test document is assigned to one of the authors from the author set using the learning model.

4.3 Evaluation Measures

The performance of the SVM in combination with various lexical features for various data set sizes and author set sizes is evaluated using accuracy and F_1 measure. The accuracy and F_1 measure are defined as follows:

Accuracy is the number of text articles from test set for which the author is correctly assigned over the total number of articles in the test set as in Eq. 1

$$\text{Accuracy} = \frac{\text{Number of documents that are correctly assigned}}{\text{Total number of test documents}} \quad (1)$$

F_1 is calculated as in Eq. 2

$$F_1 = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (2)$$

where

$$\text{precision} = \frac{\text{Number of documents correctly author assigned}}{\text{Number of documents author assigned}} \quad (3)$$

and

$$\text{recall} = \frac{\text{Number of documents correctly assigned}}{\text{Total number of test documents}} \quad (4)$$

4.4 Influence of Number of Authors in the Training Set on Authorship Attribution

In this phase, 120 text documents were considered in training set and five documents for each author in the test set. Initially, we have considered 6 authors each with 20 documents in the training set. With the help of training and testing set documents, the precision, recall, F_1 measure and accuracy were evaluated using SVM classifier with different feature vectors. The obtained results were presented in Table 1. Similarly, process is continued by considering 8, 10 and 12 authors in the training set with 15, 12 and 10 text documents for each author subsequently. From the obtained results, it is clear that as the number of authors in the training set increases, there is a significant decrease in the performance both in terms of F_1 and accuracy measures. Word unigram is outperforming compared with all other features. After the word unigram, character trigram is performing well compared with the remaining features. Character-level features performance is good when compared with word-level features.

4.5 Influence of Amount of Data per Author in the Training Set on Authorship Attribution

Articles are collected from various topics written by 12 authors. In total, we have collected 300 news articles from the websites. From the collected text documents, 240 documents are treated as training set and the remaining 60 documents are considered as test set. For the experimental evaluation, each time we have considered 5, 10, 15 and 20 documents per author. The performance of the each feature using SVM in terms of F_1 and accuracy is calculated for each data set as shown in Table 2. From the obtained results, it is clear that as the number of documents in the

Table 1 F_1 and accuracy values for number of authors with number of features using SVM for fixed data size

Feature	F_1 value						Accuracy			
	Number of authors									
	6	8	10	12	6	8	10	12		
Character unigram	0.68	0.64	0.61	0.58	0.74	0.71	0.69	0.65		
Character bigram	0.75	0.71	0.68	0.63	0.78	0.75	0.70	0.64		
Character trigram	0.82	0.78	0.75	0.69	0.85	0.81	0.74	0.71		
Character tetragram	0.79	0.75	0.76	0.67	0.82	0.79	0.77	0.74		
Word unigram	0.84	0.81	0.77	0.71	0.87	0.83	0.79	0.76		
Word bigram	0.76	0.73	0.67	0.64	0.79	0.75	0.72	0.67		
Word trigram	0.68	0.66	0.64	0.61	0.73	0.71	0.68	0.64		
Word tetragram	0.64	0.62	0.60	0.56	0.69	0.70	0.63	0.60		

Bold indicates outperforming value compared with other values

Table 2 F_1 and accuracy values for number of documents with number of features using SVM for fixed number of authors

Feature	F_1 value				Accuracy			
	Number of documents per author							
	5	10	15	20	5	10	15	20
Character unigram	0.51	0.58	0.62	0.65	0.58	0.65	0.68	0.71
Character bigram	0.55	0.63	0.65	0.71	0.61	0.64	0.70	0.75
Character trigram	0.59	0.69	0.73	0.75	0.65	0.71	0.76	0.81
Character tetragram	0.60	0.67	0.70	0.73	0.63	0.74	0.73	0.78
Word unigram	0.66	0.71	0.76	0.85	0.68	0.76	0.83	0.89
Word bigram	0.58	0.64	0.69	0.76	0.62	0.67	0.69	0.77
Word trigram	0.52	0.61	0.63	0.69	0.59	0.64	0.66	0.69
Word tetragram	0.49	0.56	0.59	0.66		0.60		

Bold indicates outperforming value compared with other values

training set increases there is significant increase in the performance both in terms of F_1 and accuracy measures. Word unigram is outperforming compared with all other features. After the word unigram, character trigram is performing well compared with the remaining features. On an average, the character-level features are exhibiting good performance compared with word-level features. The reason for good performance is, in general, character-level features will gather clues from lexical, syntactic and structural levels, and also character n-grams reduce the sparseness of the data.

5 Conclusion

Addressing the problem of authorship attribution on Telugu text is not yet attempted by any other researcher still now. The work is carried out in this paper is a real motivation towards the language. We have viewed the AA problem as a text classification problem. In this paper, it is evaluated the influence of various lexical features at character and word level with varying lengths and also empirically evaluated the impact of number of authors in the training set by keeping the total number of documents in the training set constant. Similarly, we also studied the influence of data set by keeping the number of authors in the training set as constant. In both cases, we obtained the expected results. In most cases, word unigrams and character trigrams are performing well in terms of F_1 metric and accuracy compared with other features. In this paper, for each feature vector, learning model is generated using SVM.

As a future work, we can investigate the influence of various machine learning algorithms that were investigated for generating best suitable learning models. And, it also considered various other features and its combinations to increase the performance of author identification. We can also extend the scope of study on AA by increasing the data set size.

References

1. Zhao, Y., Zobel, J.: Searching with style: authorship attribution in classic literature
2. Stamatatos, E., Fakotakis, N., Kokkinakis, G.: Automatic text categorization in terms of genre and author. *Comput. Linguist.* **26**, 471–495 (2000)
3. Holmes, D.I.: Authorship attribution. *Comput. Humanit.* **28**(2), 87–106 (1994)
4. Zhai, C.X., Lafferty, J.: Model-based feedback in the KL-divergence retrieval model. In: Proceedings of the 10th ACM CIKM International Conference on Information Knowledge Management, ACM Press, Atlanta, Georgia, USA, pp. 403–410 (2001)
5. Bozkurt, D., Baglioglu, O., Uyar, E.: Authorship attribution: performance of various features and classification methods. *Computer and information sciences* (2007)
6. Zhao, Y., Zobel, J., Vines, P.: Using relative entropy for authorship attribution. In: Proceedings of the 3rd AIRS Asian Information Retrieval Symposium, Springer, Singapore, pp. 92–105 (2006)
7. Vishnu Vardhan, B., Padmaja Rani, B., Kanaka Durga, A., Pratap Reddy, L., Vinay Babu, A.: Analysis of N-gram model on telugu document classification. In: Proceedings of 2008 IEEE Congress on Evolutionary Computation (CEC 2008), Hong Kong, pp. 3198–3202(1–6 June 2008)
8. Klinkenberg, R., Joachims, T.: Detecting concept drift with support vector machines. In: Proceedings of the 17th ICML International Conference on Machine Learning, Morgan Kaufmann Publishers, Stanford, California, USA, pp. 487–494 (2000)
9. Mosteller, F., Wallace, D.: Inference and Disputed Authorship: The Federalist. Addison-Wesley Publishing Company, USA (1964)
10. Yang, Y.M., Zhang, J., Kisiel, B.: A scalability analysis of classifiers in text categorization. In: Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM Press, Toronto, Canada, pp. 96–103 (2003)
11. Yule, G.U.: On sentence-length as a statistical characteristic of style in prose, with applications to two cases of disputed authorship. *Biometrika* **30**, 363–390 (1938)
12. Holmes, D.I.: The analysis of literary style: a review. *Roy. Stat. Soc. A* **148**(4), 328–341 (1985)
13. Baayen, H., Halteren, H.V., Neijt, A., Tweedie, F.: An experiment in authorship attribution. In: Proceedings 6th International Conference on the Statistical Analysis of Textual Data, pp. 29–37 (2002)
14. Diederich, J., Kindermann, J., Leopold, E., Paass, G.: Authorship attribution with support vector machines. *Appl. Intell.* **19**(1–2), 109–123 (2003)
15. Holmes, D.I., Robertson, M., Paez, R.: Stephen Crane and the New York tribune: a case study in traditional and non-traditional authorship attribution. *Comput. Humanit.* **35**(3), 315–331 (2001)
16. Juola, P., Baayen, H.: A controlled-corpus experiment in authorship identification by cross-entropy. *Literary Linguist. Comput.* **20**, 59–67 (2003)
17. Burrows, J.: Delta: a measure of stylistic difference and a guide to likely authorship. *Literary Linguist. Comput.* **17**, 267–287 (2002)
18. Stamatatos, E., Fakotakis, N., Kokkinakis, G.: Automatic authorship attribution. In: Proceedings of the 9th Conference of the European Chapter of the Association for Computational Linguistics, Association for Computational Linguistics, Bergen, Norway, pp. 158–164 (1999)
19. Stamatatos, E., Fakotakis, N., Kokkinakis, G.: Computer-based authorship attribution without lexical measures. *Comput. Humanit.* **35**(2), 193–214 (2001)
20. Burrows, J.: Word patterns and story shapes: the statistical analysis of narrative style. *Literary Linguist. Comput.* **2**, 61–70 (1987)
21. Binongo, J.N.G.: Who wrote the 15th book of Oz? An application of multivariate statistics to authorship attribution. *Comput. Linguist.* **16**(2), 9–17 (2003)

22. Pol, M.S.: A stylometry-based method to measure intra and inter-authorial faithfulness for forensic applications. In: SIGIR Workshop on Stylistic Analysis of Text for Information Access, ACM Press, Salvador, Bahia, Brazil (2005)
23. Kukushkina, O.V., Polikarpov, A.A., Khmelev, D.V.: Using literal and grammatical statistics for authorship attribution. *Probl. Inf. Transm.* **37**(2), 172–184 (2001)
24. Yang, Y.M., Pedersen, J.O.: A comparative study on feature selection in text categorization. In: Proceedings of the 14th ICML International Conference on Machine Learning, Morgan Kaufmann Publishers, Nashville, Tennessee, USA, pp. 412–420 (1997)
25. Farringdon, J.M.: Analysing for Authorship: A Guide to the Cusum Technique. University of Wales Press, UK (1996)
26. Benedetto, D., Caglioti, E., Loreto, V.: Language trees and zipping. *Am. Phys. Soc.* **88**(4), 048702 (2002)
27. Jurafsky, D., Martin, J.H.: Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition. Prentice Hall, USA (2000)
28. Juola, P.: What can we do with small corpora? Document categorization via cross-entropy. In: Proceedings of the Interdisciplinary Workshop on Similarity and Categorization, Edinburgh, UK (1997)
29. Kjell, B.: Authorship attribution of text samples using neural networks and bayesian classifiers. In: Proceedings of IEEE International Conference on Systems, Man and Cybernetics, IEEE Press, San Antonio, Texas, pp. 1660–1664 (1994a)
30. Domingos, P., Pazzani, M.J.: On the optimality of the simple bayesian classifier under zero-one loss. *Mach. Learn.* **29**(2/3), 103–130 (1997)
31. Gabrilovich, E., Markovitch, S.: Text categorization with many redundant features: using aggressive feature selection to make SVMs competitive with C4.5. In: Proceedings of the 21st ICML International Conference on Machine Learning, ACM Press, Banff, Alberta, Canada, pp. 321–328 (2004)
32. Zheng, R., Li, J., Chen, H., Huang, Z.: A framework for authorship identification of online messages: writing-style features and classification techniques. *J. Am. Soc. Inf. Sci. Technol.* **57**(3), 378–393 (2006)
33. Vishnu Vardhan, B., Vijaypal Reddy, P., Govardhan, A.: Corpus based extractive summarization for Indic script. In: International Conference on Asian Language Processing (IALP) IEEE computer society (IALP 2011) pp. 154–157
34. Pal Reddy, P.V., Vishnu Murthy, G., Vishnu Vardhan, B., Sarangam, K.: A comparative study on term weighting methods for automated telugu text categorization with effective classifiers. *Int. J. Data Min. Knowl. Manage. Process (IJDKP)* **3**(6) (Nov. 2013)
35. Vishnu Vardhan, B., Pal Reddy, P.V., Govardhan, A.: Analysis of BMW model for title word selection on Indic scripts. *Int. J. Comp. Appl. (IJCA)* **18**(8), 21–25 (2011)
36. Luyckx, K.: Scalability issues in authorship attribution. Ph.D thesis, Faculty of Arts and Philosophy, Dutch UPA University (2010)
37. Houvardas, J., Stamatatos, E.: N-gram feature selection for authorship identification. In: Euzennat, J., Domingue, J. (eds.) Proceeding of Artificial Intelligence: Methodology, Systems, and Applications (AIMSA), pp. 77–86. Springer, Berlin (2006)
38. Stamatatos, E.: Author identification: using text sampling to handle the class imbalance problem. *Inf. Process. Manage.* **44**(2), 790–799 (2008)

Author Index

A

- Adlin mano, R., 127
Agrawal, Samarth, 519
Agrawal, Yash, 527
Ahuja, Chakshu, 367
Akpinar, Muhammet Vefa, 559

B

- Babu, A.V., 469
Bagchi, Angshuman, 603
Balamurugan, N.B., 231, 325
Balasubramaniam, D., 127
Banerjee, Abhik, 545
Banerjee, Arundhati, 603
Banerjee, Nilotpal, 593
Banerjee, Sayan, 511
Banerjee, Suman, 239, 251
Barman, Manisha, 177
Barman, Soma, 89
Baveenther, U., 49
Bera, Rabindranath, 205, 435
Bera, Rajesh, 333
Bera, Soumyasree, 435
Bhakta, Dhananjay, 19
Bhalchandra, Parag, 1
Bharani dharan, N., 223
Bhaumik, Jaydeb, 89
Bhaumik, Shovan, 349
Biplab, Swarnakar, 59
Biswas, S., 177

C

- Chakraborty, Baisakhi, 155
Chakraborty, Souvik, 155
Chakravorty, Pragnan, 503
Chandel, Rajeevan, 527
Chatterjee, Amitava, 511
Chatterjee, Shambo, 155

- Chauhan, Seema, 251
Chellammal, N., 479
Chinnathambi, M., 223
Choudhary, Kavita, 569
Choudhury, J. Paul, 177

D

- Das, Deepanwita, 49
Das, Dhiman Kumar, 97
Das, Subhajit, 145
Das, Sudipta, 189
Das, Suman, 37
Dash, S.S., 479
Datta, Parul, 213
Datta, Sanjit Kumar, 537
Deepak, K.S., 469
Deshmukh, Nilesh, 1
Dey, Hemanta, 37
Dhiman, Rohit, 527
Dwivedi, Amit Krishna, 405

G

- Ghosh, Ambalika, 105
Ghosh, Ranjan, 37
Ghosh, Ratna, 453
Ghosh, Sumalya, 277, 287
Ghoshal, Nabin, 145, 155
Ghoshal, S.P., 545
Ghoshal, Sakti Prasad, 189, 333
Gosain, Anjana, 581
Goswami, Bhaswati, 453

H

- Hazra, Sunanda, 67
Hui, Nirmal Baran, 413

I

- Islam, Aminul, 405

J

Jaliya, U.K., 261
 Jana, Nanda Dular, 239, 251

K

Kar, Rajib, 189, 333
 Karforma, Sunil, 105
 Karthick Anand Babu, A.B., 137
 Kaushal, Sakshi, 213
 Khamitkar, Santosh, 1
 Khan, Ibrahim, 493
 Koneru, Basava Naga Girish, 357
 Kumar, Amish, 79
 Kumar, Manish, 519
 Kundu, Sumana, 11

L

Lakshmi Priya, G., 231
 Lakshminarayanan, G., 357
 Lokhande, Sakharam, 1

M

Madhu, T., 115
 Mahanti, Gautam Kumar, 197
 Maheshwari, Saurabh, 569
 Maiti, Suchismita, 79
 Maji, Bansibadan, 205, 435
 Mal, Agnish, 537
 Mal, Ashis Kumar, 277, 287, 537
 Mal, Surajit, 277, 287
 Malathy, P., 381
 Mallapur, Jayashree D., 167
 Mandal, Durbadal, 189, 333, 503
 Mandal, S., 267
 Mandal, SatyendraNath, 145
 Manikandan, S., 231, 325
 Mannapur, Ashwini V., 167
 Mantripragada, Vijaya Krishna Teja, 357
 Mekewad, Satish, 1
 Mishra, Subhash, 445
 Mishra, Vaibhav, 569
 Mitra, Anirban C., 593
 Mukherjee, V., 545

N

Nagaprasad, S., 613
 Nandi, Debashis, 79
 Nirala, C.K., 267

P

Pakala, Hara Gopal Mani, 493
 Pal, Chiranjit, 155
 Pancar, Erhan Burak, 559

Pani, Prithvi Raj, 357

Parande, S.P., 167
 Patidar, Hemant, 197
 Pitchiah, R., 423
 Premkumar, N., 307
 Priya, G. Lakshmi, 325

R

Radhakrishnan, Rahul, 349
 Raghunadha Reddy, T., 613
 Rajaram, S., 223, 307
 Raju, K.V. S.V. N., 493
 Ramaraj, N., 127
 Rana, N.K., 313
 Rathod, J.M., 261
 Ray, Arjit, 479
 Ray, Sujay, 603
 Roy, Provas Kumar, 67
 Roy, Shibendu Shekhar, 299
 Roy, Somnath, 519

S

Saha, Debasree, 239
 Saha, Manika, 453
 Sahoo, Sukadev, 97
 Sailaja, M., 115
 Samanta, Jagannath, 89
 Sarangi, Somnath, 393
 Saraswathi, D., 325
 Sarker, Goutam, 11, 19
 Seelam, Koteswararao, 115
 Selvi, C., 367
 Sharma, Ganga, 581
 Sharma, Pratima, 413
 Shaw, Binod, 545
 Shunmugalatha, A., 381
 Siddheshwar, Sen, 59
 Singh, Abhinoy Kumar, 349
 Singh, Rajendra Prasad, 445
 Sinha, Anubhav, 405
 Sivakumar, R., 137
 Sivasankar, E., 367, 423
 Sur, Samarendra Nath, 205, 435

T

Thaineesh, P., 381
 Theodore Chandra, S., 231
 Tomar, Nutan Kumar, 349

U

Uma, Datta, 59
 Umapathi, N., 127
 Upadhyay, Ekta M., 313

V

- Vanitha, P., [231](#)
Varaprasad, P.L. H., [493](#)
Venkatasubramani, V.R., [307](#)
Verma, Alok, [393](#)
Verma, Bhupendra, [27](#)
Vignesh, K., [307](#)
Vignesh Raja, P., [423](#)

- Vijayopal Reddy, P., [613](#)
Vinaya Babu, A., [613](#)
VishnuVardhan, B., [613](#)

Y

- Yadav, Ram Narayan, [445](#)
Yadav, Shailendra Singh, [27](#)