

Lecture Notes in Networks and Systems 94

Mohan L. Kolhe  
Shailesh Tiwari  
Munesh C. Trivedi  
Krishn K. Mishra *Editors*

# Advances in Data and Information Sciences

Proceedings of ICDIS 2019

 Springer

# Lecture Notes in Networks and Systems

Volume 94

## Series Editor

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

## Advisory Editors

Fernando Gomide, Department of Computer Engineering and Automation—DCA,  
School of Electrical and Computer Engineering—FEEC, University of Campinas—  
UNICAMP, São Paulo, Brazil

Okyay Kaynak, Department of Electrical and Electronic Engineering,  
Bogazici University, Istanbul, Turkey

Derong Liu, Department of Electrical and Computer Engineering, University  
of Illinois at Chicago, Chicago, USA; Institute of Automation, Chinese Academy  
of Sciences, Beijing, China

Witold Pedrycz, Department of Electrical and Computer Engineering,  
University of Alberta, Alberta, Canada; Systems Research Institute,  
Polish Academy of Sciences, Warsaw, Poland

Marios M. Polycarpou, Department of Electrical and Computer Engineering,  
KIOS Research Center for Intelligent Systems and Networks, University of Cyprus,  
Nicosia, Cyprus

Imre J. Rudas, Óbuda University, Budapest, Hungary

Jun Wang, Department of Computer Science, City University of Hong Kong,  
Kowloon, Hong Kong

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

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

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

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

**\*\* Indexing: The books of this series are submitted to ISI Proceedings, SCOPUS, Google Scholar and Springerlink \*\***

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

Mohan L. Kolhe · Shailesh Tiwari ·  
Munesh C. Trivedi · Krishn K. Mishra  
Editors

# Advances in Data and Information Sciences

Proceedings of ICDIS 2019

 Springer

*Editors*

Mohan L. Kolhe  
Smart Grid and Renewable Energy  
University of Agder  
Kristiansand, Norway

Munesh C. Trivedi  
Department of Computer Science  
and Engineering  
NIT Agartala  
Tripura, India

Shailesh Tiwari  
Department of Computer Science  
and Engineering  
ABES Engineering College  
Ghaziabad, Uttar Pradesh, India

Krishn K. Mishra  
Computer Science and Engineering  
Motilal Nehru National Institute  
of Technology Allahabad  
Prayagraj, India

ISSN 2367-3370

ISSN 2367-3389 (electronic)

Lecture Notes in Networks and Systems

ISBN 978-981-15-0693-2

ISBN 978-981-15-0694-9 (eBook)

<https://doi.org/10.1007/978-981-15-0694-9>

© Springer Nature Singapore Pte Ltd. 2020

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

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

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

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

# **Organizing Committee**

## **Chief Patron**

Raja Anirudh Pal Singh, Vice President, Balwant Educational Society, Agra, India

## **Patron**

Yuvraj Ambreesh Pal Singh, Secretary, Balwant Educational Society, Agra, India

## **General Chair**

Prof. (Dr.) B. S. Kushwaha, Director, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

## **General Co-chair**

Prof. (Dr.) Pankaj Gupta, Director (F&A), Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

## **Conference Chair**

Prof. (Dr.) Brajesh Kumar Singh, Head of Department, Computer Science and Engineering, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

## **Publication Chairs**

Prof. (Dr.) Apoorva Behari Lal, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

Er. Anshul Kumar Singh, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

Er. Lavkush Sharma, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

Er. Aalisha Goel, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

## **Publicity and Social Media Chairs**

Dr. Vivek Srivastav, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

Er. Aman Singh, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

Dr. Amit Kumar Singh, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

Er. Saumya Tripathi, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

## **Workshop Chairs**

Prof. (Dr.) Sapna Tomar, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

Er. Ashok Kumar, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

Dr. Anurag Kulshreshtha, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

Er. Prachi Pundhir, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

## **Posters and Demo Chairs**

Prof. (Dr.) Shraddha Rani Singh, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

Dr. D. S. Tomar, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

Dr. Dushyant Singh, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

Er. Alok Singh Jadaun, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

## **Sponsorship and Exhibits Chairs**

Dr. Sachipati Pandey, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

Er. Jay Kumar, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

Er. Amit Agarwal, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India

Er. Geetanjali Singh, Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India



# Preface

The ICDIS is a major multidisciplinary conference organized with the objective of bringing together researchers, developers, and practitioners from academia and industry working in all areas of computer and computational sciences. It is organized specifically to help computer industry to derive the advances of next-generation computer and communication technology. Researchers invited to speak will present the latest developments and technical solutions.

Technological developments all over the world are dependent upon globalization of various research activities. Exchange of information, innovative ideas is necessary to accelerate the development of technology. Keeping this ideology in preference, the Second International Conference on Data and Information Sciences (ICDIS 2019) has been organized at Raja Balwant Singh Engineering Technical Campus, Bichpuri, Agra, India, during March 29–30, 2019.

The Second International Conference on Data and Information Sciences has been organized with a foreseen objective of enhancing the research activities at a large scale. Technical Program Committee and Advisory Board of ICDIS 2019 include eminent academicians, researchers, and practitioners from abroad as well as from all over the nation.

In ICDIS 2019 proceedings, selected manuscripts have been subdivided into five tracks named: Advanced Communications and Security, Intelligent Computing Techniques, Intelligent Hardware and Software Design, Web and Informatics, and Intelligent Image Processing. A sincere effort has been made to make it an immense source of knowledge by including 61 manuscripts in this proceedings volume. The selected manuscripts have gone through a rigorous review process and are revised by authors after incorporating the suggestions of the reviewers.

ICDIS 2019 received more than 250 submissions from around 550 authors of different countries such as India, Tanzania, China, Malaysia, Bangladesh, Sri Lanka, and many more. Each submission has been gone through the plagiarism check. On the basis of plagiarism report, each submission was rigorously reviewed by at least two reviewers. Even some submissions have more than two reviews. On the basis of these reviews, 61 high-quality papers were selected for publication in two proceedings volumes, with an acceptance rate of 24.6%.

We are thankful to the keynote speakers Prof. Erma Suryani, ITS, Indonesia, and Prof. K. K. Biswas, IIT Delhi, India, for enlightening the participants with their knowledge and insights. We are also thankful for delegates and the authors for their participation and their interest in ICDIS 2019 as a platform to share their ideas and innovation. We are also thankful to Prof. Dr. Janusz Kacprzyk, Series Editor, LNNS, Springer Nature, and Mr. Aninda Bose, Senior Editor, Springer Nature, India, for providing guidance and support. Also, we extend our heartfelt gratitude to the reviewers and technical program committee members for showing their concern and efforts in the review process. We are indeed thankful to everyone directly or indirectly associated with the conference organizing team for leading it toward the success.

Although utmost care has been taken in compilation and editing, a few errors may still remain. We request the participants to bear with such errors and lapses (if any). We wish you all the best...

Agra, India

Mohan L. Kolhe  
Shailesh Tiwari  
Munesh C. Trivedi  
Krishn K. Mishra

# Contents

## Advanced Communications and Security

<b>Weighted Dissemination of Bundles in Probabilistic Spray and Wait Routing Protocol</b> . . . . .	3
Diksha Sharma, Sanjay Kumar and Naresh Kumar Nagwani	
<b>Study of Network-Induced Delays on Networked Control Systems</b> . . . .	13
Jitendra Kumar, Vishal Goyal and Devbrat Gupta	
<b>Text-to-Image Encryption and Decryption Using Piece Wise Linear Chaotic Maps</b> . . . . .	23
K. Abhimanyu Kumar Patro, Shashwat Soni, V. K. Sharma and Bibhudendra Acharya	
<b>Security Threats, Attacks, and Possible Countermeasures in Internet of Things</b> . . . . .	35
Shams Tabrez Siddiqui, Shadab Alam, Riaz Ahmad and Mohammed Shuaib	
<b>Securing IoT-Driven Remote Healthcare Data Through Blockchain</b> . . . . .	47
Sarthak Gupta, Virain Malhotra and Shailendra Narayan Singh	
<b>A Review of Big Data Challenges and Preserving Privacy in Big Data</b> . . . . .	57
Anil Sharma, Gurwinder Singh and Shabnum Rehman	
<b>Dual-Layer DNA-Encoding–Decoding Operation Based Image Encryption Using One-Dimensional Chaotic Map</b> . . . . .	67
K. Abhimanyu Kumar Patro, M. Prasanth Jagapathi Babu, K. Pavan Kumar and Bibhudendra Acharya	

**Simple Permutation and Diffusion Operation Based Image Encryption Using Various One-Dimensional Chaotic Maps: A Comparative Analysis on Security** . . . . . 81  
 Dasari Sravanthi, K. Abhimanyu Kumar Patro, Bibhudendra Acharya and M. Prasanth Jagapathi Babu

**Integration of Wireless Sensor Networks with Cloud Towards Efficient Management in IoT: A Review** . . . . . 97  
 Rajendra Kumar Dwivedi, Nikita Kumari and Rakesh Kumar

**Reliability-Based Resource Scheduling Approach Using Hybrid PSO-GA in Mobile Computational Grid** . . . . . 109  
 Krishan Veer Singh and Zahid Raza

**Internet of Things (IoT) Enabling Technologies, Requirements, and Security Challenges** . . . . . 119  
 Shadab Alam, Shams Tabrez Siddiqui, Ausaf Ahmad, Riaz Ahmad and Mohammed Shuaib

**Impact of Network Load for Anomaly Detection in Software-Defined Networking** . . . . . 127  
 Ashish Gupta, Bharat Didwania, Gaurav Singh, Hari Prabhat Gupta, Rahul Mishra and Tanima Dutta

**An Extended Playfair Encryption Technique Based on Fibonacci Series** . . . . . 135  
 Mohd Vasim Ahamad, Mohd Imran, Nazish Siddiqui and Tasleem Jamal

**An Automated System for Epileptic Seizure Detection Using EEG** . . . . 147  
 Bilal Alam Khan, Anam Hashmi and Omar Farooq

**Addressing Security and Privacy Issues of Load Balancing Using Hybrid Algorithm** . . . . . 157  
 T. Subha

**Key Management Scheme for Secure Group Communication** . . . . . 171  
 Om Pal and Bashir Alam

**Lightweight Hardware Architecture for Eight-Sided Fortress Cipher in FPGA** . . . . . 179  
 Nivedita Shrivastava and Bibhudendra Acharya

**Two-Dimensional Hybrid Authentication for ATM Transactions** . . . . . 191  
 M. F. Mridha, Jahir Ibna Rafiq and Wahid Uz Zaman

**Intelligent Computing Techniques**

**Artificial Neural Network Based Load Balancing in Cloud Environment** ..... 203  
 Sarita Negi, Neelam Panwar, Kunwar Singh Vaisla and Man Mohan Singh Rauthan

**Maximum Power Point Tracking Using a Hybrid Fuzzy Logic Control** ..... 217  
 Amruta S. Deshpande and Sanjaykumar L. Patil

**Differential Evolution Algorithm Using Enhance-Based Adaption Mutant Vector** ..... 227  
 Shailendra Pratap Singh and Deepak Kumar Singh

**Standard Library Tool Set for Rough Set Theory on FPGA** ..... 237  
 Vanita Agarwal and Rajendrakumar A. Patil

**A Comparison of the Effectiveness of Two Novel Clustering-Based Heuristics for the  $p$ -Centre Problem** ..... 247  
 Mahima Yadav and V. Prem Prakash

**Half-Life Teaching Factor Based TLBO Algorithm** ..... 257  
 Ruchi Mishra, Nirmala Sharma and Harish Sharma

**Multilingual Data Analysis to Classify Sentiment Analysis for Tweets Using NLP and Classification Algorithm** ..... 271  
 Pragati Goel, Vikas Goel and Amit Kumar Gupta

**Intelligent Hardware and Software Design**

**Pedestrian–Autonomous Vehicles Interaction Challenges: A Survey and a Solution to Pedestrian Intent Identification** ..... 283  
 Pranav Pandey and Jagannath V. Aghav

**Code Profiling Analysis of Rough Set Theory on DSP and Embedded Processors for IoT Application** ..... 293  
 Vanita Agarwal, Rajendrakumar A. Patil and Jyoti Adwani

**Design and Analysis of IoT-Based System for Crowd Density Estimation Techniques** ..... 307  
 Ajitesh Kumar and Mona Kumari

**Video-Transmission-Based Condition Monitoring of Solar Panels Using QR Code** ..... 317  
 Akash Singh Chaudhary, Isha and D. K. Chaturvedi

**Effects of Activation Function and Input Function of ANN for Solar Power Forecasting** ..... 329  
 Isha, Akash Singh Chaudhary and D. K. Chaturvedi

<b>An Integrated Approach Toward Smart Parking Implementation for Smart Cities in India</b> .....	343
Ishan Kumar, Prashant Manuja, Yashpal Soni and Narendra Singh Yadav	
<b>Distributed Processes Scheduling Based on Evolutionary Approach</b> .....	351
Santosh Kumar, Gaurav Dubey and Shailesh Tiwari	
<b>Self-driving Cars: An Overview of Various Autonomous Driving Systems</b> .....	361
V. Shreyas, Skanda N. Bharadwaj, S. Srinidhi, K. U. Ankith and A. B. Rajendra	
<b>Internet of Things: Industry Use Cases (SAP-HCP)</b> .....	373
Avaneesh Kumar Vats and Nagsen Wankhede	
<b>Web and Informatics</b>	
<b>Organizational Readiness for Managing Large-Scale Data Storage in Virtualized Server Environments</b> .....	381
Said Ally	
<b>Classification of Forest Cover Type Using Random Forests Algorithm</b> .....	395
Arvind Kumar and Nishant Sinha	
<b>Impact of Noisy Labels in Learning Techniques: A Survey</b> .....	403
Nitika Nigam, Tanima Dutta and Hari Prabhat Gupta	
<b>Performance Analysis of Schema Design Approaches for Migration from RDBMS to NoSQL Databases</b> .....	413
Basant Namdeo and Ugrasen Suman	
<b>A Time Delay Neural Network Acoustic Modeling for Hindi Speech Recognition</b> .....	425
Ankit Kumar and R. K. Aggarwal	
<b>A Review on Offensive Language Detection</b> .....	433
Rahul Pradhan, Ankur Chaturvedi, Aprna Tripathi and Dilip Kumar Sharma	
<b>Attribute-Based Elliptic Curve Encryption for Security in Sensor Cloud</b> .....	441
Munish Saran, Rajendra Kumar Dwivedi and Rakesh Kumar	
<b>Predictive Model Prototype for the Diagnosis of Breast Cancer Using Big Data Technology</b> .....	455
Ankita Sinha, Bhaswati Sahoo, Siddharth Swarup Rautaray and Manjusha Pandey	

<b>Recent Dimensions of Data Science: A Survey</b> .....	465
Sinkon Nayak, Mahendra Kumar Gourisaria, Manjusha Pandey and Siddharth Swarup Rautaray	
<b>Sentiment Analysis: Usage of Text and Emoji for Expressing Sentiments</b> .....	477
Shelley Gupta, Archana Singh and Jayanthi Ranjan	
<b>Factors Affecting Psychological State of Youth in India</b> .....	487
Jagreeti Kaur, Archana Singh, Sumit Kumar and Sunil Kumar	
<b>Enhancing Personalized Response to Product Queries Using Product Reviews Incorporating Semantic Information</b> .....	497
Payal Aich, Manju Venugopalan and Deepa Gupta	
<b>Enhancing Future Relationship in Social Network Using Semantics Prediction to Predict Links</b> .....	511
Snigdha Luthra, Gursimran Kaur and Dilbag Singh	
<b>Privacy Rights for Digital Assets and Digital Legacy Right for Posterity: A Survey</b> .....	521
Amit Sudan, Munish Sabharwal, Wan Khairuzzaman Wan Ismail and Yogesh Kumar	
<b>Intelligent Image Processing</b>	
<b>Ear Detection and Recognition Techniques: A Comparative Review</b> .....	533
Pallavi Srivastava, Diwakar Agrawal and Atul Bansal	
<b>Automatic Detection of Sleep Spindles Using Time Domain Features</b> .....	545
Ghania Fatima, Omar Farooq and Shikha Singh	
<b>A Review on Lung and Nodule Segmentation Techniques</b> .....	555
Bhawana Kamble, Satya Prakash Sahu and Rajesh Doriya	
<b>Wavelet Decomposition Based Authentication Scheme for Dental CBCT Images</b> .....	567
Ashish Khatter, Nitya Reddy and Anita Thakur	
<b>A Comparative Analysis of Different Violence Detection Algorithms from Videos</b> .....	577
Piyush Vashistha, Juginder Pal Singh and Mohd Aamir Khan	
<b>Minimizing Synchronization Error in Compressed Domain Watermarking</b> .....	591
Tanima Dutta, Aishwarya Soni and Hari Prabhat Gupta	

<b>Deep Learning Architectures for Computer Vision Applications: A Study</b> .....	601
Randheer Bagi, Tanima Dutta and Hari Prabhat Gupta	
<b>Robust Reversible Watermarking for Grayscale Medical Images</b> .....	613
Tanima Dutta, Randheer Bagi and Hari Prabhat Gupta	
<b>Improved Detection of Kidney Stone in Ultrasound Images Using Segmentation Techniques</b> .....	623
Rati Goel and Anmol Jain	
<b>Non-adaptive and Adaptive Filtering Techniques for Fingerprint Pores Extraction</b> .....	643
Diwakar Agarwal and Atul Bansal	
<b>Character and Mesh Optimization of Modern 3D Video Games</b> .....	655
Ragib Hasan, Sumittra Chakraborti, Md. Zonieed Hossain, Taukir Ahamed, Md. Abdul Hamid and M. F. Mridha	
<b>Image Watermarking Scheme Using Cuckoo Search Algorithm</b> .....	667
Gaurav Dubey, Charu Agarwal, Santosh Kumar and Harivansh Pratap Singh	
<b>A Survey of Latent Fingerprint Indexing and Segmentation Based Matching</b> .....	677
Harivans Pratap Singh and Priti Dimri	
<b>Author Index</b> .....	687



# Editors and Contributors

## About the Editors

**Prof. Mohan L. Kolhe** is a professor of electrical power engineering with a focus on smart grids and renewable energy at the Faculty of Engineering and Science, University of Agder (Norway). He has more than twenty-five years of international academic experience in electrical and renewable energy systems. He is a leading renewable energy technologist and has previously held academic positions at prestigious universities around the globe, including University College London (UK / Australia), University of Dundee (UK); University of Jyväskylä (Finland); and the Hydrogen Research Institute, QC (Canada).

**Prof. Shailesh Tiwari** currently works as a professor at the Computer Science and Engineering Department, ABES Engineering College, Ghaziabad, India. He is an alumnus of Motilal Nehru National Institute of Technology, Allahabad, India. His primary areas of research are software testing, implementation of optimization algorithms, and machine learning techniques in various problems. He has published more than 50 papers in leading international journals and conference proceedings, and serves as an editor for various Scopus, SCI, and E-SCI-indexed journals. He has organized several international conferences under the banner of the IEEE and Springer. He is a senior member of the IEEE, member of the IEEE Computer Society, and a Fellow of the Institution of Engineers (FIE).

**Prof. Munesh C. Trivedi** currently works as a professor at the Computer Science and Engineering Department, ABES Engineering College, Ghaziabad, India. He has published 20 textbooks and 80 research papers in various leading international journals and conference proceedings. He has received Young Scientist and numerous other awards from national and international forums, and has organized several international conferences sponsored by the IEEE, ACM, and Springer. He serves on the review panel of the IEEE Computer Society, the International Journal of Network Security, Pattern Recognition Letters, and Computer & Education, and

as an executive committee member of the IEEE UP Section, IEEE India Council, and IEEE Asia Pacific Region 10.

**Prof. Krishn K. Mishra** currently works as a visiting faculty at the Department of Mathematics & Computer Science, University of Missouri, St. Louis, USA. He is an alumnus of Motilal Nehru National Institute of Technology, Allahabad, India, which is also his home working institute. His primary areas of research include evolutionary algorithms, optimization techniques, and the design and analysis of algorithms. He has published more than 50 papers in international journals and international conference proceedings. He has served as a program committee member for several conferences and also edited Scopus and SCI-indexed journals. He has 15 years of teaching and research experience.

## Contributors

**Md. Abdul Hamid** Department of Computer Science and Engineering, University of Asia Pacific, Dhaka, Bangladesh

**K. Abhimanyu Kumar Patro** Department of Electronics and Telecommunication Engineering, National Institute of Technology Raipur, Raipur, India

**Bibhudendra Acharya** Department of Electronics and Telecommunication Engineering, National Institute of Technology Raipur, Raipur, India

**Jyoti Adwani** Centre for VLSI and Nanotechnology, VNIT, Nagpur, Maharashtra, India

**Charu Agarwal** Department of Computer Science and Engineering, Ajay Kumar Garg Engineering College, Ghaziabad, India

**Diwakar Agarwal** Electronics & Communication Engineering, GLA University, Mathura, UP, India

**Vanita Agarwal** Electronics and Telecommunication Department, College of Engineering Pune, Pune, Maharashtra, India

**R. K. Aggarwal** Department of Computer Engineering, National Institute of Technology, Kurukshetra, Haryana, India

**Jagannath V. Aghav** College of Engineering, Wellesley Rd, Shivajinagar, Pune, Maharashtra, India

**Diwakar Agrawal** GLA University, Mathura, India

**Mohd Vasim Ahamad** Department of Computer Science & Engineering, University Polytechnic, Integral University, Lucknow, India

**Taukir Ahamed** Department of Computer Science and Engineering, University of Asia Pacific, Dhaka, Bangladesh

**Ausaf Ahmad** Department of Computer Science, Aligarh Muslim University, Aligarh, India

**Riaz Ahmad** Department of Computer Science, Aligarh Muslim University, Aligarh, India

**Payal Aich** Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Bengaluru, India

**Bashir Alam** Faculty of Engineering & Technology, Department of Computer Engineering, Jamia Millia Islamia, New Delhi, India

**Shadab Alam** Department of Computer Science, Jazan University, Jizan, Saudi Arabia

**Said Ally** The Open University of Tanzania, Dar es Salaam, Tanzania

**K. U. Ankith** Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India

**Randheer Bagi** Department of Computer Science and Engineering, IIT (BHU) Varanasi, Varanasi, India

**Atul Bansal** Electronics & Communication Engineering, GLA University, Mathura, UP, India

**Skanda N. Bharadwaj** Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India

**Sumittra Chakraborti** Department of Computer Science and Engineering, University of Asia Pacific, Dhaka, Bangladesh

**Ankur Chaturvedi** GLA University, Mathura, India

**D. K. Chaturvedi** Department of Electrical Engineering, Faculty of Engineering, Dayalbagh Educational Institute, Agra, India;  
Faculty of Engineering, Department of Electrical Engineering, Dayalbagh Educational Institute, Agra, India

**Akash Singh Chaudhary** Department of Electrical Engineering, Faculty of Engineering, Dayalbagh Educational Institute, Agra, India;  
Faculty of Engineering, Department of Electrical Engineering, Dayalbagh Educational Institute, Agra, India

**Amruta S. Deshpande** College of Engineering Pune, Pune, India

**Bharat Didwania** Department of Electrical Engineering, IIT (BHU), Varanasi, India

**Priti Dimri** Department of Computer Science and Applications, G.B. Pant Engineering College, Ghurdauri, Uttarakhand, India

**Rajesh Doriya** Department of Information Technology, National Institute of Technology, Raipur (C.G), India

**Gaurav Dubey** Department of Computer Science and Engineering, ABES Engineering College, Ghaziabad, India

**Tanima Dutta** Department of Computer Science and Engineering, IIT (BHU) Varanasi, Varanasi, India

**Rajendra Kumar Dwivedi** Department of Computer Science and Engineering, Madan Mohan Malaviya University of Technology, Gorakhpur, UP, India

**Omar Farooq** Department of Electronics Engineering, Aligarh Muslim University, Aligarh, UP, India

**Ghania Fatima** Department of Electronics Engineering, Aligarh Muslim University, Aligarh, India

**Pragati Goel** Shri Venketeshwara University, Gajraula, UP, India

**Rati Goel** ABES Engg. College, Ghaziabad, India

**Vikas Goel** Ajay Kumar Garg Engineering College, Ghaziabad, UP, India

**Mahendra Kumar Gourisaria** KIIT Deemed to be University, Bhubaneswar, India

**Vishal Goyal** Department of Electronics and Communication Engineering, GLA University, Mathura, UP, India

**Amit Kumar Gupta** KIET Group of Institutions, Ghaziabad, UP, India

**Ashish Gupta** Department of Computer Science and Engineering, IIT (BHU), Varanasi, India

**Deepa Gupta** Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Bengaluru, India

**Devbrat Gupta** Department of Electronics and Communication Engineering, GLA University, Mathura, UP, India

**Hari Prabhat Gupta** Department of Computer Science and Engineering, IIT (BHU) Varanasi, Varanasi, India

**Sarthak Gupta** ASET, Amity University, Noida, India

**Shelley Gupta** Department of Information Technology, ABES Engineering College, Ghaziabad, Uttar Pradesh, India

**Ragib Hasan** Department of Computer Science and Engineering, University of Asia Pacific, Dhaka, Bangladesh

**Anam Hashmi** Department of Electronics Engineering, Aligarh Muslim University, Aligarh, UP, India

**Mohd Imran** Department of Computer Engineering, ZHCET, Aligarh Muslim University, Aligarh, India

**Isha** Faculty of Engineering, Department of Electrical Engineering, Dayalbagh Educational Institute, Agra, India

**Wan Khairuzzaman Wan Ismail Sulaiman** AL Rajhi School of Business Albukavarivah, Al Bukayriyah, Saudi Arabia

**Anmol Jain** ABES Engg. College, Ghaziabad, India

**Tasleem Jamal** Department of Information Technology, REC, Azamgarh, India

**Bhawana Kamble** Department of Information Technology, National Institute of Technology, Raipur (C.G), India

**Gursimran Kaur** Department Apex Institute of Technology, Chandigarh University, Ajitgarh, India

**Jagreeti Kaur** ABES Engineering College, Ghaziabad, India

**Bilal Alam Khan** Department of Electronics Engineering, Aligarh Muslim University, Aligarh, UP, India

**Mohd Aamir Khan** GLA University, Mathura, India

**Ashish Khatter** Department of Electronics and Communication Engineering, Amity University, Noida, India

**Ajitesh Kumar** GLA University, Mathura, UP, India

**Ankit Kumar** Department of Computer Engineering, National Institute of Technology, Kurukshetra, Haryana, India;  
Computer Science and Engineering Department, Galgotias University, Greater Noida, India

**Arvind Kumar** Bennett University, Greater Noida, India

**Ishan Kumar** Manipal University Jaipur, Jaipur, India

**Jitendra Kumar** Department of Electronics and Communication Engineering, GLA University, Mathura, UP, India

**Rakesh Kumar** Department of Computer Science and Engineering, Madan Mohan Malaviya University of Technology, Gorakhpur, UP, India

**Sanjay Kumar** Department of Information Technology, National Institute of Technology, Raipur, India

**Santosh Kumar** Department of Computer Science and Engineering, ABES Engineering College, Ghaziabad, India

- Sumit Kumar** Amity University, Noida, Uttar Pradesh, India
- Sunil Kumar** Amity University, Noida, Uttar Pradesh, India
- Yogesh Kumar** Department of Computer Science and Engineering, Chandigarh University, Mohali, Punjab, India
- Mona Kumari** GLA University, Mathura, UP, India
- Nikita Kumari** Department of Computer Science and Engineering, Madan Mohan Malaviya University of Technology, Gorakhpur, UP, India
- Snigdha Luthra** Department Apex Institute of Technology, Chandigarh University, Ajitgarh, India
- Virain Malhotra** ASET, Amity University, Noida, India
- Prashant Manuja** Manipal University Jaipur, Jaipur, India
- Rahul Mishra** Department of Computer Science and Engineering, IIT (BHU), Varanasi, India
- Ruchi Mishra** Rajasthan Technical University, Kota, India
- M. F. Mridha** Department of Computer Science and Engineering, University of Asia Pacific, Dhaka, Bangladesh
- Naresh Kumar Nagwani** Department of Computer Science and Engineering, National Institute of Technology, Raipur, India
- Basant Namdeo** International Institute of Professional Studies, DAVV, Indore, India
- Sinkon Nayak** KIIT Deemed to be University, Bhubaneswar, India
- Sarita Negi** Uttarakhand Technical University, Dehradun, Uttarakhand, India; SOET, HNBGU, Srinagar, Garhwal, Uttarakhand, India
- Nitika Nigam** Department of Computer Science and Engineering, IIT (BHU), Varanasi, India
- Om Pal** Ministry of Electronics and Information Technology Government of India, New Delhi, India; Faculty of Engineering & Technology, Department of Computer Engineering, Jamia Millia Islamia, New Delhi, India
- Manjusha Pandey** KIIT Deemed University, Bhubaneswar, India
- Pranav Pandey** College of Engineering, Wellesley Rd, Shivajinagar, Pune, Maharashtra, India
- Neelam Panwar** SOET, HNBGU, Srinagar, Garhwal, Uttarakhand, India
- Rajendrakumar A. Patil** Electronics and Telecommunication Department, College of Engineering Pune, Pune, Maharashtra, India

**Sanjaykumar L. Patil** College of Engineering Pune, Pune, India

**K. Abhimanyu Kumar Patro** Department of Electronics and Telecommunication Engineering, National Institute of Technology Raipur, Raipur, India

**K. Pavan Kumar** Department of Electronics and Telecommunication Engineering, National Institute of Technology Raipur, Raipur, India

**Rahul Pradhan** GLA University, Mathura, India

**M. Prasanth Jagapathi Babu** Department of Electronics and Telecommunication Engineering, National Institute of Technology Raipur, Raipur, India

**V. Prem Prakash** Department of Electrical Engineering, Faculty of Engineering, Dayalbagh Educational Institute (Deemed University), Dayalbagh, Agra, India

**Jahir Ibna Rafiq** University of Asia Pacific, Dhaka, India

**A. B. Rajendra** Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India

**Jayanthi Ranjan** Department of Information Technology, Institute of Management Technology, Ghaziabad, India

**Siddharth Swarup Rautaray** KIIT Deemed University, Bhuneshwar, India

**Man Mohan Singh Rauthan** SOET, HNBGU, Srinagar, Garhwal, Uttarakhand, India

**Zahid Raza** School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi, India

**Nitya Reddy** Department of Electronics and Communication Engineering, Amity University, Noida, India

**Shabnum Rehman** School of Computer Applications, Lovely Professional University, Phagwara, Punjab, India

**Munish Sabharwal** Department of Computer Science and Engineering, Chandigarh University, Mohali, Punjab, India

**Bhaswati Sahoo** KIIT Deemed University, Bhuneshwar, India

**Satya Prakash Sahu** Department of Information Technology, National Institute of Technology, Raipur (C.G), India

**Munish Saran** Department of CSE, MMMUT Gorakhpur, Gorakhpur, India

**Diksha Sharma** Department of Information Technology, National Institute of Technology, Raipur, India

**Dilip Kumar Sharma** GLA University, Mathura, India

**Harish Sharma** Rajasthan Technical University, Kota, India

**Nirmala Sharma** Rajasthan Technical University, Kota, India

**V. K. Sharma** Department of Electronics and Telecommunication Engineering, National Institute of Technology Raipur, Raipur, India

**Anil Sharma** School of Computer Applications, Lovely Professional University, Phagwara, Punjab, India

**V. Shreyas** Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India

**Nivedita Shrivastava** Department of Electronics and Telecommunication Engineering, National Institute of Technology Raipur, Raipur, India

**Mohammed Shuaib** Department of Computer Science, Jazan University, Jizan, Saudi Arabia

**Nazish Siddiqui** Department of Computer Science & Engineering, University Polytechnic, Integral University, Lucknow, India

**Shams Tabrez Siddiqui** Department of Computer Science, Jazan University, Jizan, Saudi Arabia

**Archana Singh** Department of Information Technology, Amity School of Engineering and Technology, Noida, India;  
Amity University, Noida, Uttar Pradesh, India

**Deepak Kumar Singh** Department of Computer Science and Engineering, Sachedeva Institute of Technology, Frah, Mathura, UP, India

**Dilbag Singh** Department Apex Institute of Technology, Chandigarh University, Ajitgarh, India

**Gaurav Singh** Department of Electrical Engineering, IIT (BHU), Varanasi, India

**Gurwinder Singh** School of Computer Applications, Lovely Professional University, Phagwara, Punjab, India

**Harivansh Pratap Singh** Department of Computer Science and Engineering, ABES Engineering College, Ghaziabad, India

**Juginder Pal Singh** GLA University, Mathura, India

**Krishan Veer Singh** School of Computer and Systems Sciences, Jawaharlal Nehru University, New Delhi, India

**Shailendra Narayan Singh** ASET, Amity University, Noida, India

**Shailendra Pratap Singh** Department of Computer Science and Engineering, Bundelkhand Institute of Engineering and Technology Jhansi, Jhansi, UP, India

**Shikha Singh** Department of Electronics Engineering, Aligarh Muslim University, Aligarh, India



**Ankita Sinha** KIIT Deemed University, Bhuneshwar, India

**Nishant Sinha** Pitney Bowes Software, Noida, India

**Aishwarya Soni** Department of Computer Science and Engineering, IIT (BHU) Varanasi, Varanasi, India

**Shashwat Soni** Department of Electronics and Telecommunication Engineering, National Institute of Technology Raipur, Raipur, India

**Yashpal Soni** Manipal University Jaipur, Jaipur, India

**Dasari Sravanthi** Department of Information Technology, National Institute of Technology Raipur, Raipur, India

**S. Srinidhi** Department of Information Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India

**Pallavi Srivastava** GLA University, Mathura, India

**T. Subha** Sri Sai Ram Engineering College, Chennai, India

**Amit Sudan** Department of Computer Science and Engineering, Chandigarh University, Mohali, Punjab, India

**Ugrasen Suman** School of Computer Science and IT, DAVV, Indore, India

**Anita Thakur** Department of Electronics and Communication Engineering, Amity University, Noida, India

**Shailesh Tiwari** Department of Computer Science and Engineering, ABES Engineering College, Ghaziabad, India

**Aprna Tripathi** GLA University, Mathura, India

**Kunwar Singh Vaisla** B.T.KIT, Dwarhat, Uttarakhand, India

**Piyush Vashistha** GLA University, Mathura, India

**Avaneesh Kumar Vats** Energy Efficiency Services Ltd, EESL, Delhi, India

**Manju Venugopalan** Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Bengaluru, India

**Nagsen Wankhede** Energy Efficiency Services Ltd, EESL, Delhi, India

**Mahima Yadav** Department of Electrical Engineering, Faculty of Engineering, Dayalbagh Educational Institute (Deemed University), Dayalbagh, Agra, India

**Narendra Singh Yadav** Manipal University Jaipur, Jaipur, India

**Wahid Uz Zaman** University of Asia Pacific, Dhaka, India

**Md. Zoniceed Hossain** Department of Computer Science and Engineering, University of Asia Pacific, Dhaka, Bangladesh

# **Advanced Communications and Security**

# Weighted Dissemination of Bundles in Probabilistic Spray and Wait Routing Protocol



Diksha Sharma, Sanjay Kumar and Naresh Kumar Nagwani

**Abstract** Delay Tolerant Network is a new emerging technology, delivering messages in a challenged network termed as Intermittently Connected Networks (ICNs), lacking continuous end-to-end connectivity, having low data rate and high propagation delay. Routing of bundles is an area of interest in DTN. Spray and Wait is a DTN routing protocol that outstrips other DTN routing protocols ProPHET, Epidemic in performance metric overhead ratio. The performance of the Spray and Wait protocol in other metrics is intended to be elevated in this work. The proposed algorithm implements weighted dissemination of messages instead of even dissemination in the spraying phase. Number of replicas to be transmitted to an encountered node is decided on the basis of its delivery probability. The proposed algorithm transmits less number of replicas to the node having greater delivery probability as they have more chances of encountering the destination node. The algorithm explores more possible ways to find a suitable hop as it also considers giving packets to nodes having low probability considering the situation that it might encounter a node having the best probability to deliver.

**Keywords** ICNs · Delay Tolerant Network · Probabilistic Spray and Wait · Uneven distribution

---

D. Sharma (✉) · S. Kumar  
Department of Information Technology, National Institute of Technology, Raipur, India  
e-mail: [dsharma.mtech2017.it@nitrr.ac.in](mailto:dsharma.mtech2017.it@nitrr.ac.in)

S. Kumar  
e-mail: [skumar.it@nitrr.ac.in](mailto:skumar.it@nitrr.ac.in)

N. K. Nagwani  
Department of Computer Science and Engineering, National Institute of Technology,  
Raipur, India  
e-mail: [nknagwani.cs@nitrr.ac.in](mailto:nknagwani.cs@nitrr.ac.in)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_1](https://doi.org/10.1007/978-981-15-0694-9_1)

# 1 Introduction

Interplanetary Communication project was launched with the purpose of establishing data communication in deep space environment. Conventional communication protocols prove to be inefficient in implementing communication in deep space environment. TCP/IP is the conventional protocol designed for the networks with a guaranteed end-to-end path between the source and receiver. The pair of communication devices in deep space is not to be connected by a guaranteed and reliable end-to-end path and also have very high propagation delay among them. The TCP/IP protocol tends to fail in the given network scenario. These networks were given a proper classification as challenged networks or Intermittently Connected Networks which share the common characteristics of unreliable and no continuous end-to-end path, high propagation delay, and low data rates. Underwater transmission, wildlife tracking and deep space communication are a few networks qualifying the tag of ICNs. Achieving transmission in these challenged networks was possible by the introduction of a new transfer principle “Store and Forward”.

Delay Tolerant Networking [1] with acronym DTN was first introduced by Kevin Fall in the year 2003 in a conference. It is a networking approach designed to tackle the challenges in a disrupted and disconnected network failing to deliver end-to-end connections. The initial motivation to design DTN was transmitting data over profound distance as in the case of space communication or even on an interplanetary scale. Such environment has inexorable long latency ranging from hours to, sometimes, days and year. Delay tolerant networking is a networking approach that is designed to enable communication when data is being transmitted and received between millions of miles of propagation delay. This results in the delay of data at the receiver end. The Disruption Tolerant Networking approach ensures reliable communication in intermittent networks by overcoming the potential disruption that occurred in the process due to the low data rate and no continuous path. Delay/Disruption Tolerant Network is a solution to providing reliable communication in a challenged network which has significant delays and disruptions.

The term bundle was introduced in the architecture of DTN. The bundle layer is the end-to-end message overlay which exists above the transport layer and under the applications of the network. DTN nodes are the devices which implement these bundle layers. The DTN nodes carrying bundles to be forwarded have to follow a route pattern to reach for the destination node. For connected networks with constant end-to-end connectivity, routing protocols are proposed. These routing protocols cannot be implemented in Intermittently Connected Networks where a fixed path is not guaranteed. Therefore, need for designing a DTN-based routing protocol was felt to facilitate the transmission of bundles in a harsh environment.

Researchers have implemented various routing protocols for the DTN environment. Section 2 of this work describes the work done so far in the main routing protocols in DTN. It explains Epidemic, Spray and Wait, and ProPHET routing protocols. Section 3 of this work proposes a new approach of the Spray and Wait routing protocol that disseminates a different number of copies to the encountered node

considering its capability to deliver. The algorithm implementing the weighted-based dissemination is described in the same section. Section 4 presents the simulation environment and the results of the above-proposed model when implemented in the ONE simulator. The work is concluded with Sect. 5 briefing the work of this paper and also listing the future work intended to be carried afterward.

## 2 Related Work

The Delay Tolerant Network [1] proposed by Kevin Fall initially meant to provide message transmission in a space environment as part of the Interplanetary Project. Intermittently Connected Networks (ICNs) with no reliable end-to-end connection are not the usual networks for the traditional Internet protocol for transmission. DTN uses the principle of hop to hop or “store carry forward” to transmit bundles in challenged environment. The DakNet project [2] was an application of Delay Tolerant Network which was implemented with the intention of providing Internet data to people living in rural areas where proper Internet infrastructure is not established. Low-cost kiosks were installed in the areas equipped with a wifi adapter. Kiosk controller controls the device where people can access government documents or other required document. Public transport with an adapter and storage is the one which synchronizes the kiosk. It carries the data of the kiosk to the nearby Internet provider and uploads the content to the Internet. Other applications of Delay Tolerant Network are proposed and listed in [3].

Routing protocols in a Delay Tolerant Network are cataloged in a survey [4–6] which defines Epidemic, Spray and Wait, and ProPHET routing protocols with their pros and cons. Performance of all three routing protocols is compared on the basis of varying buffer size of DTN nodes involved in routing and varying message or bundle size to be transmitted. Results of the comparisons are tabulated which serve as a base to implement improvisation in the routing protocols. “Epidemic” [7] protocol uses a general way of delivering bundles. The DTN node carrying the bundle transfers one copy of the bundle to any other node it comes in contact with. Selection of the next node is random in the Epidemic protocol. The protocol provides high delivery probability as it enumerates all possible paths to the destination. The number of copies generated in the process is very high in number. It results in overuse of network resources thus increasing the overhead ratio. It follows blind selection of the next node encountered and does not follow any criteria for rejecting a node. Reference [8] discuss the enhanced version of Epidemic in the energy saving perspective.

The Spray and Wait routing protocol [9] was proposed which restricts the flooding of bundles to a limit. It floods blindly to the neighboring node but with constraints of the number of nodes and the number of copies of the bundle. It reduces the network overhead compared to Epidemic, as the number of copies is restricted by an equation which generates “L” number of copies to relay.

ProPHET [10] Probabilistic routing protocol for Intermittently Connected Networks is an information-based flooding proposed to remove blind dissemination of bundles to nodes. Working process of the protocol starts from the source node

intending to transmit a bundle to a destination node. The nodes contain a vector displaying their delivery probability to other nodes. The host node on coming in contact with the other node, exchanges its probability vector. The host node verifies if the other node's delivery probability is superior to it, if true then the host node transmits a copy of the bundle to that node. It repeats flooding of the bundle to every encountered node having delivery probability greater than itself. However in its initial step, it utilizes uniform probability distribution which deteriorates the overall performance of the protocol compared to Epidemic.

Papers have been published proposing improvisation in the existing protocols to exploit their cons and outperform the base protocol. Pi, a practical incentive protocol for Delay Tolerant Networks [11] was proposed to deal with selfish nodes in DTN. Probability based the Spray and Wait protocol in Delay Tolerant Networks [12] is an improvisation of the Spray and Wait protocol which utilizes ProPHET information-based routing to select the next node. The author has tabulated the results showing the comparison of the proposed Spray and Wait and the base protocols. The proposed model according to the author when simulated has performed better in delivery ratio and average delay compared to Spray and Wait.

Routing protocols for DTN can be simulated in any simulator like OMNET++ and NS2 Java-based simulator. But these simulators do not provide environment scenarios and tools to efficiently implement a simulation. Performance of the DTN-based simulation is dependent on the characteristics of the DTN node involved and the movement pattern. The above simulators are not decorated with the required simulation tools for implementing DTN-based protocols or scenarios. Opportunistic Network Environment [13] (ONE) simulator was proposed, dedicated to ease the implementation of new routing protocols and scenarios for Delay Tolerant Network applications. It has powerful tools including predefined movement models—Random Walk, Random Way Point, and Shortest Path Map Based. Routing protocols Epidemic, ProPHET [10], ProPHETV2 [14], MaxProp [15], and Spray and Wait are implemented in the simulator.

The principle binary Spray and Wait routing algorithm transfers equal amounts of replicas in a single phase. The distribution algorithm assigns  $L/2$  replicas to the encountered node irrespective of their capability to deliver the message. The improvised version Probabilistic Spray and Wait discovers less number of paths as it ignores the path traversed by the node having delivery probability less than the carrier node. The proposed algorithm makes use of the delivery probability of the encountered node. The node having good chances to encounter the destination node is assigned less replicas.

### **3 Weighted Dissemination of Bundles in Probabilistic Spray and Wait**

The proposed Spray and Wait implements weighted approach to decide on the number of copies of the bundle to be transferred to encountered node. The delivery probability implemented in ProPHET is utilized to estimate the probability of encountering

the destination node. This delivery probability acts as a decision factor in estimating the number of replications to be transferred to the next hop. This does not completely eliminate the nodes having low probability than the host node, such that low probability node may have contact with other nodes having much more delivery probability than the host node.

Thus the proposed method explores more route possibility. Let node X have generated L copies of the bundle and intend to transmit it to node Z. It encounters a neighboring node Y having  $P(Y, Z)$  less than  $P(X, Z)$  of node X, then it does not ignore the node Y considering that the route followed by node Y may cross path with a node say A such that  $P(A, Z) \gg P(X, Z)$ . The node having higher delivery probability implicates that it frequently encounters the destination node, thus the nodes having higher probability are given less number of replications.

### 3.1 Proposed Algorithm

#### 3.1.1 Weighted Dissemination Based Spray Phase

The total number of replicas to be transmitted to the encountered node is calculated by considering the delivery probability of that node for a given bundle. Equation 1 generates the number of copies (Fig. 1).

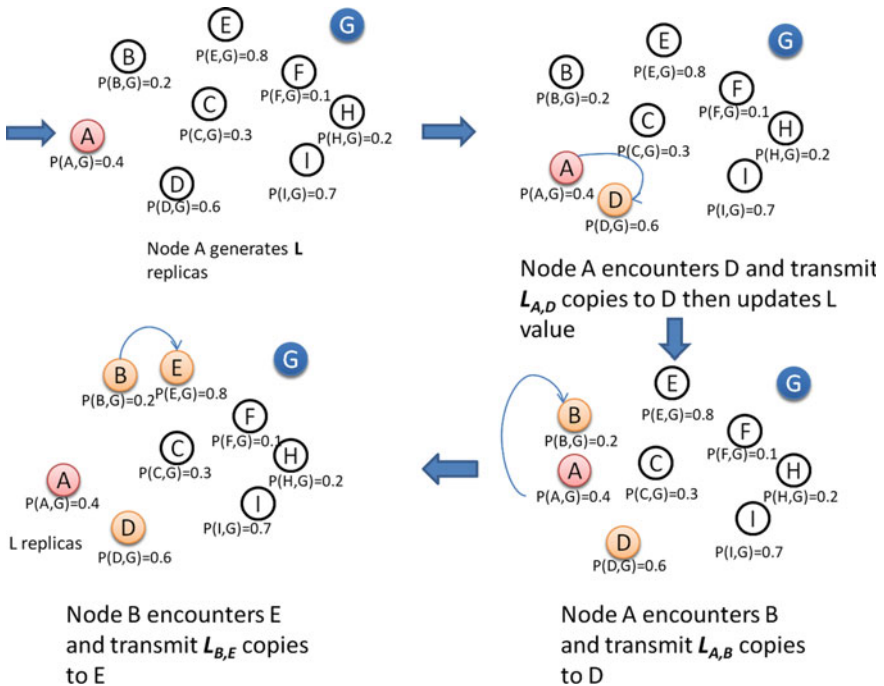


Fig. 1 Weighted dissemination of bundles in Probabilistic Spray and Wait routing protocol

$$copies(L_{A,C}) = \lceil L_A[1 - P(C, B)]/3 \rceil \quad (1)$$

$L_{A,C}$  = Number of copies of a bundle to be transferred from node A to C.

$L_A$  = Copies of a bundle generated in node A.

Algorithm 1 depicts the proposed algorithm that distributes weighted number of replicas to the encountered node.

---

### Algorithm 1 Spray Phase

---

```

1: Node X generates  $L_x$  replicas of a message with destination Node Z
2: Generate Delivery Probabilities of Nodes.
3: DTN Node X encounters DTN Node Y
4: while node encountered do
5:   X and Y exchange summary vectors
6:   if bundle then
7:     if FirstBundle(bundle,Y) then
8:       if Node Y is Destination then
9:         transfeBundle(X,Y)
10:      else
11:         $D_x = \text{DeliveryProbability}(\text{bundle}, X)$ 
12:         $D_y = \text{DeliveryProbability}(\text{bundle}, Y)$ 
13:         $\text{replica} = \text{GetReplicas}(\text{bundle}, X)$ 
14:        if  $\text{replica} > 1$  then
15:          Calculate replicas to be transferred

```

$$copies(L_{A,C}) = \lceil L_A[1 - P(C, B)]/3 \rceil \quad (2)$$

```

16:      Transfer(X,Y,copies)
17:      SetReplicas = replica - copies
18:    else
19:      Node X enter wait state for bundle
20:    end if
21:  end if
22: end if
23: end if
24: end while

```

---

## 4 Simulations and Results

ONE (Opportunistic Network Environment) simulator is the simulator dedicated especially for developing and simulating applications and routing protocols for Delay Tolerant Environment. Powerful simulations tools facilitating simulation are available in the ONE simulator. DTN routing protocols are judged on the basis of performance matrices. These metrics contribute in implementing an effective routing protocol for Delay Tolerant Network. These metrics are Delivery Probability, Average Latency and Overhead Ratio.

Simulation time was set for 10 h with the word size of  $4500 \times 3400$ . Buffer size of the DTN nodes is 20MB and the size of the message is set between the range of



500kb–1 MB and the Time to Live of the message is set to 300 min. DTN nodes are transmitting with the speed of 2 Mbps. The simulation is tested for the sparse and dense networks. In sparse network, the number of nodes are set to 50 and in the dense network, the number of nodes are 100.

The proposed model is compared with the Probabilistic Spray and Wait routing protocol. The sparse network involves very less number of nodes; thus, the proposed model considers all the nodes as a potential forwarder to explore more number of paths from the source to destination. The model considers the weight of the encountered node so as to calculate the efficient number of replicas to be forwarded. Table 1 tabulates the delivery probability obtained after the simulation which proves that the model is more suitable for the sparse network with less number of DTN nodes.

The proposed protocol when ran for a dense network with 200 nodes in the same simulation environment yields better delivery probability than Probabilistic Spray and Wait. Table 2 tabulates the statistics of the total number of messages created in both scenarios for routing protocols Probabilistic Spray and Wait and weighted dissemination based Probabilistic Spray and Wait. In VDTN, the position of the DTN node is not fixed and may change at any moment. It is tedious to predict the position of every DTN node. Thus deciding the forwarder depending only on the delivery probability (ProPHET) may not be efficient in terms of the number of paths covered. The algorithm does not reject any node; it assumes every node as a potential forwarder which allows to cover as many paths as possible. In addition, it restricts the number of replicas to be forwarded by considering the delivery probability of the node (Fig. 2).

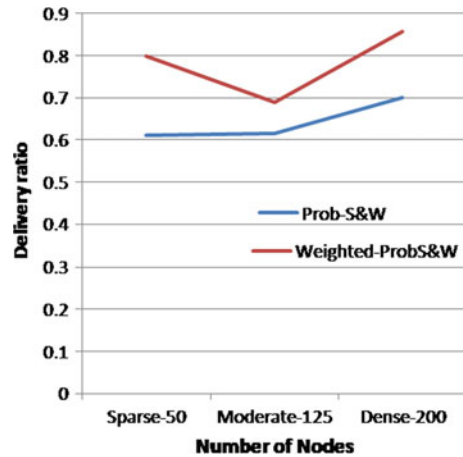
**Table 1** Delivery ratio

Network (Number of nodes)	Probabilistic Spray and Wait	Weighted Dissemination based Spray and Wait
Sparse—50	0.61	0.70
Moderate—125	0.615	0.69
Dense—200	0.8	0.857

**Table 2** Relayed message statistics

Network (Number of nodes)	Message created	Probabilistic Spray and Wait	Weighted Dissemination based Spray and Wait	Percentage increased
Sparse—50	590	359	413	15.0
Moderate—125	590	362	407	12.43
Dense—200	610	488	524	7.4

**Fig. 2** Graph depicts the variance in delivery ratio with respect to network model Sparse, Moderate, and Dense. Proposed weighted dissemination based model performs better in sparse network with less nodes when compared to Probabilistic Spray and Wait



## 5 Conclusion

DTN routing protocols provide a means to transmit bundles in ICNs. These are updated and improvised with respect to the environment and different scenarios. In this work, we have proposed uneven or weighted distribution of bundles in Probabilistic Spray and Wait which calculates the total number of replicas to be transmitted based on its delivery probability. Probabilistic Spray and Wait ignores nodes with low delivery probability which may encounter another node with best delivery probability in its path. The proposed model weighs the encountered node and calculates suitable number of replicas to be transferred. Simulation results depict that the proposed model outperforms the Probabilistic Spray and Wait protocol in terms of delivery ratio. Average delay for bundle has to be low for any routing protocol, thus, further, we intend to reduce the average delay of the proposed routing protocol.

## References

1. Fall, K. (2003). A delay-tolerant network architecture for challenged internets. In *Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*. ACM.
2. Pentland, A., Fletcher, R., & Hasson, A. (2004). Daknet: Rethinking connectivity in developing nations. *Computer*, 37(1), 78–83.
3. Gao, L., Yu, S., Luan, T. H., & Zhou, W. (2015). *Delay tolerant networks and their applications*. New York: Springer.
4. Khabbaz, M. J., Assi, C. M., & Fawaz, W. F. (2012). Disruption-tolerant networking: A comprehensive survey on recent developments and persisting challenges. *IEEE Communications Surveys & Tutorials*, 14(2), 607–640.

5. Liu, M., Yan, Y., & Qin, Z. (2011). A survey of routing protocols and simulations in delay-tolerant networks. In *International Conference on Wireless Algorithms, Systems, and Applications*. Berlin: Springer.
6. Jones, E. P. C., et al. (2007). Practical routing in delay-tolerant networks. *IEEE Transactions on Mobile Computing*, 6(8), 943–959.
7. Vahdat, A., & Becker, D. (2000). Epidemic routing for partially connected ad hoc networks.
8. De Rango, F., Amelio, S., & Fazio, P. (2013). Enhancements of epidemic routing in delay tolerant networks from an energy perspective. In *2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC)*. IEEE.
9. Spyropoulos, T., Psounis, K., & Raghavendra, C. S. (2005). Spray and wait: An efficient routing scheme for intermittently connected mobile networks. In *Proceedings of the 2005 ACM SIGCOMM Workshop on Delay-Tolerant Networking*. ACM.
10. Lindgren, A., et al. (2012). Probabilistic routing protocol for intermittently connected networks. No. RFC 6693.
11. Lu, R., et al. (2010). Pi: A practical incentive protocol for delay tolerant networks. *IEEE Transactions on Wireless Communications*, 9(4).
12. Kim, E.-H., et al. (2014). Probability-based spray and wait protocol in delay tolerant networks. In *International Conference on Information Networking (ICOIN)*. IEEE.
13. Kern, A., Ott, J., & Kken, T. (2009). The ONE simulator for DTN protocol evaluation. In *Proceedings of the 2nd International Conference on Simulation Tools And Techniques*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).
14. Grasic, S., et al. (2011). The evolution of a DTN routing protocol-PRoPHETv2. In *Proceedings of the 6th ACM workshop on Challenged Networks*. ACM.
15. Burgess, J., et al. (2006). Maxprop: Routing for vehicle-based disruption-tolerant networks. In *INFOCOM 2006: 25th IEEE International Conference on Computer Communications*. IEEE.

# Study of Network-Induced Delays on Networked Control Systems



Jitendra Kumar, Vishal Goyal and Devbrat Gupta

**Abstract** Networked control systems (NCSs) are significant and foremost multi-disciplinary research areas for many decades. This paper is mainly oriented toward recent developments and challenges of network-induced delays due to inclusion of data network in NCSs. Network delays deteriorate the control performance and stability of the NCSs. The time-varying delays can be measured in real time by calculating the time difference of sending and receiving control packets. Various compensation techniques are reviewed to mitigate the effect of constant, time-varying, and stochastic delay. Lastly, some conclusions are drawn and the future research scope is directed.

**Keywords** Networked control systems · Network-induced delays · Communication networks · Study · Delay compensation

## 1 Introduction

The classical definition of Networked control systems (NCSs) can be given as: When a generalized feedback control system is closed via a communication channel, which may be connected with other nodes externally to the control system, then it is called an NCS. In another way, NCSs are the basic feedback control systems in which the control loops are connected via a real-time communication network [1–4]. The simplified model of NCS is shown in Fig. 1. Literature suggest that NCSs have been one of the leading investigated topics in multidisciplinary area for many decades. In classical feedback systems, the interconnection among the sensors, actuators, and

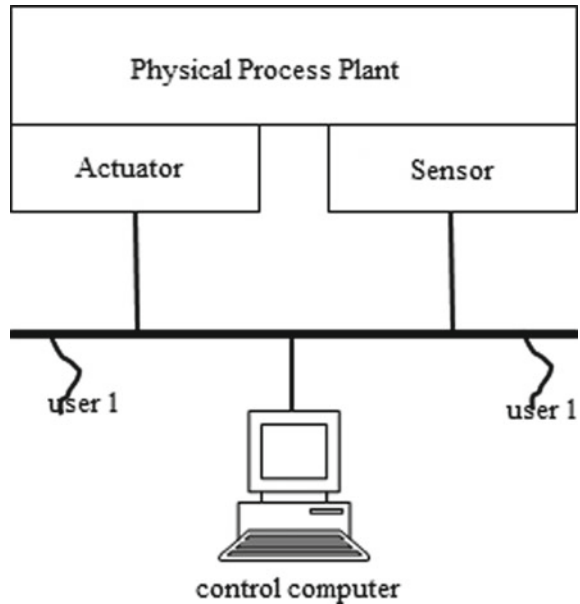
---

J. Kumar · V. Goyal · D. Gupta (✉)  
Department of Electronics and Communication Engineering, GLA University, Mathura 281406,  
UP, India  
e-mail: [devbrat.gupta1@gmail.com](mailto:devbrat.gupta1@gmail.com)

J. Kumar  
e-mail: [jitendra.kumar@gla.ac.in](mailto:jitendra.kumar@gla.ac.in)

V. Goyal  
e-mail: [vishal.glaitm@gmail.com](mailto:vishal.glaitm@gmail.com)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_2](https://doi.org/10.1007/978-981-15-0694-9_2)

**Fig. 1** Simplified NCSs

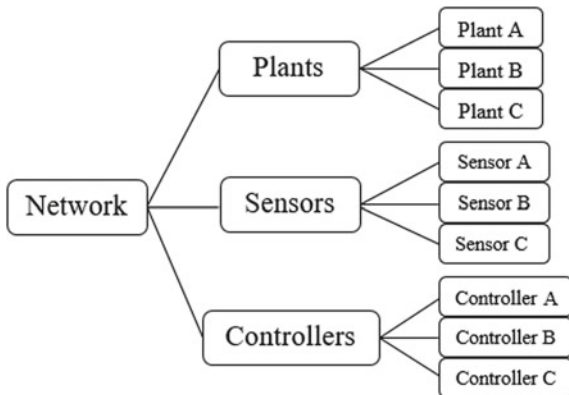
controllers are mainly realized by port-to-port connections by the communication network medium, due to which it creates many difficulties such as wiring complexities, low flexibility, and maintenance. These types of difficulties appear everywhere which motivated the rise of networked control systems with lesser cabling costs, cheaper hardware network, modularity, and flexibility in system design where automated controlled plants with increasing complexity are situated [5].

NCSs have increasing popularity among various industries with their variety of advantages over traditional control systems. It deals with interference, network congestion, data dropout, efficient data communication. It combines major industrial fields like automation and control field, computer science and network field, etc. [2].

In modern control systems, the goal is to design stable and feasible control techniques that can achieve various tasks with lesser reconfiguration cost and with ease of maintenance. A point-to-point framework is inactive from a reconfigurable opinion and it does not address the issues like interchangeability and reliability [6, 7]. The induction of communication networks gave the idea of remotely controlling a system which arises NCS. In recent years, use of wireless communication medium has been revolutionized. A conceptual model of NCS where information such as reference input, plant output, controlled output, etc., is shared via a communication network among control system elements such as actuators to drive control outputs as seen in Fig. 2 [1].

Basically, NCSs are divided into two classes: the first one is control-over-network and second is control of network. The study of performance issues of network such as networking protocol, routing control, the efficacy of the control system with network as data transmission medium (either shared or wirelessly) is being achieved in control

**Fig. 2** Conceptual model of NCS



of network. This paper represents the control-over-network part. The research of networked control system is focused on two types of quality of the system, i.e., the quality of service (QoS) and the quality of control (QoC). In the present study, it has been focused on QoC to review the stability performance of the system due to effect of network-induced delays furthermore [3].

## 2 NCSs: Brief Evolution

Origins of control systems can be noticeable before 1868 when famous physicist J. C. Maxwell conducted dynamics analysis of the centrifugal governor [8]. In the late 1950s, the computer was added in control system and further it has been extensively used. Honeywell proposed TDC-2000 in mid 1980s [9, 10]. The enhancement of shared data networks such as slotted ALOHA widely evolved to use modern network protocols [4]. In early years, Gupta and Chow surveyed the history and evolution of NCSs. In 1969, the Advanced Research Projects Agency situated in U.S. (Department of Defense) has established the first working data switching network, i.e., ARPANET. Around 1980s, controller area network (CAN) was introduced as one of the fieldbus. The time-sensitive decentralized control, i.e., fieldbus was used in industrial network around 1988. Nowadays, wireless NCSs (WNCSs) are advent and standard for high-level data rates and integrity [1]. The motive behind WNCS is ease of installation, modularity, and rapid deployments of copious fascinating applications such as smart highways (or bridges), factory automation, vehicle communication, etc. Sensor nodes are commercially accessible in abundance due to brisk development in communication, sensing hardware, and low-power computational ability. Military, teleoperation, telerobotics, and medical applications can often use optical fiber network to guarantee speedy, robust, and unflinching communication. The Internet of Things (IoT) is suggested in many applications due to low cost where the plant and the controller are remotely located. Almost all the available networks have

some small delay which subsequently deteriorates the control performance. While overviewing the analysis and synthesis of NCSs, delays are the main concern toward getting better stability performance. So certain aspects of network-induced delays are considered [10–15].

### 3 Literature Review on Various Delays

The inclusion of data networks in feedback loops of NCSs, give many advantages such as cost-effectiveness, robustness, and flexibility toward applications used. But some issues like network delays arise due to the insertion of these data networks and they degrade the NCS control performances which also affect the stability of system. Tipsuwan and Chow suggest some fundamental and recent control methodologies for NCSs to overcome some of the challenges and issues [4]. Yang has suggested a basic architecture and reviewed two important categories related to conventional large-scale NCSs [5]. The first one is control analysis and synthesis, the other one is network scheduling, protocol, and architecture. Brindha and Mendiratta have surveyed the development history of NCS furthermore and suggested the improvement of NCS performance in areas like propagation time and reduction of overall network-induced time lag with synthesis of robust and optimal controller [2]. Xia et al. summarized various phases in NCSs like quantizer, estimator, fault detection, and network predictive control and also proposed the future advancement in cloud techniques and cloud control systems [3]. This paper is mainly focused on quantization and how quantization improves the stability performance of the system.

Zhang et al. made an extensive survey about recent advancements in NCSs in his research work [7]. The synthesis, analysis, and modeling of the system have been majorly focused in his work. A new delay compensation algorithm with a feedback control law is proposed, which are connected via the CAN buses addition to a time-domain Smith predictor. It was observed that if the resulting delays are too large for an NCS with time constraints, the performance of the NCS will be degraded. This could eventually lead to possible physical harm to the controlled process or even threaten human lives. For example, in traffic NCS problem control-over-network can be observed [16–20].

Further, the concern of stability for discrete-time time-delay systems is reexamined. The Lyapunov–Krasovskii method is the most common technique to analyze the stability of time-delayed systems. To check the stability of the system, it has to find the maximum bounded delayed region such that time-delay system remains stable within this region [19]. Short time-varying delay is proposed within the region of maximal bound for defined stabilization techniques. On the basis of these techniques, stability criteria are defined. When NCS is utilized with STVD, it gets converted into time-varying discrete-time system with the help of robust control methods changed into corresponding time-invariant system. Zhang et al. proposed a delay reduction technique in which NCSs are linked via CAN where a feedback control law is suggested with the help of time-domain smith predictor to estimate the future state [14].

The asymptotic stability property of a feedback system by using augmentation is analyzed and obtained bilinear matrix inequality is transformed into a linear matrix inequality. Schenato explained the framework for optimal estimation design with two different time-invariant estimation architecture which does not depend upon the packet delay [20]. But tradeoffs appeared between packet loss and packet delay because the sensor measurements and their control packets are matter of random delay and loss simultaneously.

Shi and Yu focused on the output feedback stabilization of NCSs where the random delays are framed as Markov chain to design of a controller which is based on both the delays, i.e., measurement-to-controller and controller-to-final control element. After that closed-loop control system was transformed and output feedback controller was explored through linear matrix inequality to support stochastic stabilization [21]. Zhang et al. [14] proposed an output feedback delay control technique in which the stochastic network induced delay is modeled as Markov chain. A framework as Markovian jump linear system was designed for closed-loop control system for controller design. Heemels et al. show the tradeoffs between network delays, transmission interval, and their stability performance. Modeling of sampled-data NCS with random delays, packet loss, and quantization are performed as a nonlinear time-delay system with two consecutive delay mechanisms. Further, LK functional is used to solve the network-based  $H_\infty$  control problems in [22–28].

## 4 Challenges in Control-Over-Network

### 4.1 Issues on Control Performances

When network delays and data collisions are not taken into consideration, the control algorithm is simple and stable but problem persists when random time-varying delay occurs [29].

### 4.2 Limitations on Networks

Due to the insertion of the network and usage of a communication channel in NCSs, it degrades the stability as well as performance of the system. Such problems are appeared due to signal sampling, data quantization, communication delays (network-induced random delays), packet loss, packet disorder, channel fading, medium access constraints, power constraints, network constraints, etc. [2].



### 4.3 Delays

In general, delays are two types one is constant delays and other is random delays [18]. At the beginning of NCS-related research, the characteristics modeling or random delays are difficult so the easier approach is to model the delay as a constant directly. Random (Stochastic) delays are divided into two parts; one is dependent as their probability distribution governed by Markov chain while other is mutually independent stochastic delay. Random delays are trickier time-varying systems. NCSs are stable for all constant delays but become unstable when the delay varies [30–32].

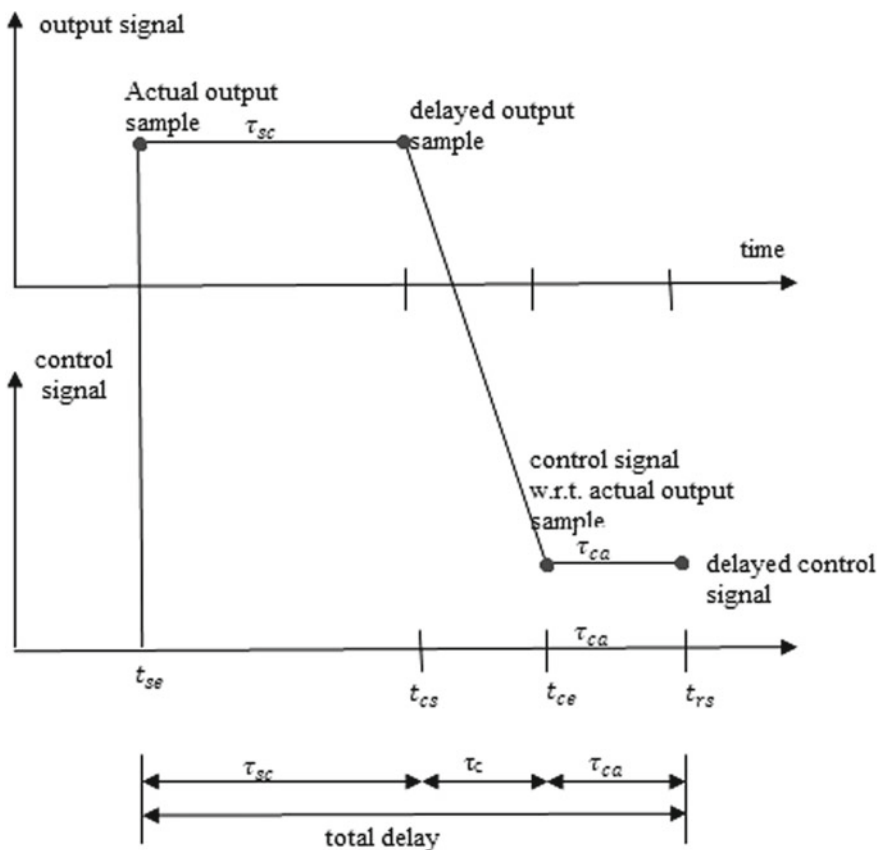


Fig. 3 Diagram of network-induced delays [4]

#### 4.4 Network-Induced Delays

When communication networks are induced in NCSs data transfer between the controller and a system installed remotely, it shows network delays [33, 34]. Tipsuwan and Chow suggested the direction of data transfers as there are many delays like measuring device-to-controller delay  $\tau_{sc}$  and the controller-to-final control element delay  $\tau_{ca}$  including computational delay are calculated in [4] as,

$$\tau_{sc} = t_{cs} - t_{se} \quad (1)$$

$$\tau_{ca} = t_{rs} - t_{ce} \quad (2)$$

where  $t_{cs}$  defined as the time taken for the controller to collect the data of measurement from sensor,  $t_{se}$  is defined as the time taken from sensor to error signal,  $t_{rs}$  is the time taken from reference signal to sensor and  $t_{ce}$  is the time taken for the primary controller to send the signal from transducer. A network delay for NCS formulations is shown in Fig. 3. The three basic delays, i.e., propagation delay, frame time delay and waiting time delay occur on a LAN i.e., local area network. Propagations of delays in NCSs are shown in Fig. 3.

### 5 Conclusion and Future Scope

NCS is multidisciplinary research area with broad applications. It uses data network to connect the components of the control system. This arises many technical issues like time delay, packet dropout, bandwidth limitation, quantization, etc. Some of the recent advancements in delay compensation techniques and modeling are exclusively summarized. However, this review does not consider other significant issues like medium access constraint, power constraints, sampling, and channel fading. Although NCSs have been very propitious research fields for many years, there are alarming and unresolved issues like explicit dependency between network delays and packet dropouts to be considered for future research.

### References

1. Gupta, R. A., & Chow, M. Y. (2010). Networked control system: Overview and research trends. *IEEE Transactions on Industrial Electronics*, 57(7), 2527–2535.
2. Brindha, M., & Mendiratta, J. K. (2013). Networked control system—A survey. *International Journal of Modern Education and Computer Science*, 5(6), 42.
3. Xia, Y. Q., Gao, Y. L., Yan, L. P., & Fu, M. Y. (2015). Recent progress in networked control systems—A survey. *International Journal of Automation and Computing*, 12(4), 343–367.

4. Tipsuwan, Y., & Chow, M. Y. (2003). Control methodologies in networked control systems. *Control Engineering Practice*, 11(10), 1099–1111.
5. Ulz, T., Pieber, T., Steger, C., Maticsek, R., & Bock, H. (2017, September). Towards trustworthy data in networked control systems: A hardware-based approach. In *2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation* (pp. 1–8). IEEE.
6. Yang, T. C. (2006). Networked control system: A brief survey. *IEEE Proceedings-Control Theory and Applications*, 153(4), 403–412.
7. Zhang, D., Shi, P., Wang, Q. G., & Yu, L. (2017). Analysis and synthesis of networked control systems: A survey of recent advances and challenges. *ISA Transactions*, 66, 376–392.
8. Maxwell, J. C. (1868). I. On governors. *Proceedings of the Royal Society of London*, 16, 270–283.
9. Guo, L., Tang, Y., Liu, Z., & Xiong, W. (2010, June). The theory and architecture of network control system. In *2010 International Conference on Intelligent Computing and Cognitive Informatics* (pp. 183–186). IEEE.
10. Goodwin, G. C., Haimovich, H., Quevedo, D. E., & Welsh, J. S. (2004). A moving horizon approach to networked control system design. *IEEE Transactions on Automatic Control*, 49(9), 1427–1445.
11. Ke-You, Y., & Li-Hua, X. (2013). Survey of recent progress in networked control systems. *Acta Automatica Sinica*, 39(2), 101–117.
12. Zhang, X. M., Han, Q. L., & Yu, X. (2016). Survey on recent advances in networked control systems. *IEEE Transactions on Industrial Informatics*, 12(5), 1740–1752.
13. Richard, J. P. (2003). Time-delay systems: An overview of some recent advances and open problems. *Automatica*, 39(10), 1667–1694.
14. Zhang, H., Zhang, Z., Wang, Z., & Shan, Q. (2016). New results on stability and stabilization of networked control systems with short time-varying delay. *IEEE Transactions on Cybernetics*, 46(12), 2772–2781.
15. Hespanha, J. P., Naghshtabrizi, P., & Xu, Y. (2007). A survey of recent results in networked control systems. *Proceedings of the IEEE*, 95(1), 138–162.
16. Liu, S., Liu, P. X., & Wang, X. (2017). Stability analysis and compensation of network-induced delays in communication-based power system control: A survey. *ISA Transactions*, 66, 143–153.
17. Gao, H., Meng, X., & Chen, T. (2008). Stabilization of networked control systems with a new delay characterization. *IEEE Transactions on Automatic Control*, 53(9), 2142–2148.
18. Ge, Y., Chen, Q., Jiang, M., & Huang, Y. (2013). Modeling of random delays in networked control systems. *Journal of Control Science and Engineering*, 2013, 8.
19. Zhang, H., Shi, Y., Wang, J., & Chen, H. (2018). A new delay-compensation scheme for networked control systems in controller area networks. *IEEE Transactions on Industrial Electronics*.
20. Schenato, L. (2006, December). Optimal estimation in networked control systems subject to random delay and packet loss. In *Proceedings of the 45th IEEE Conference on Decision and Control* (pp. 5615–5620). IEEE.
21. Shi, Y., & Yu, B. (2009). Output feedback stabilization of networked control systems with random delays modeled by Markov chains. *IEEE Transactions on Automatic Control*, 54(7), 1668–1674.
22. Zhang, J., Lam, J., & Xia, Y. (2014). Output feedback delay compensation control for networked control systems with random delays. *Information Sciences*, 265, 154–166.
23. Heemels, W. M. H., Teel, A. R., Van de Wouw, N., & Nesić, D. (2010). Networked control systems with communication constraints: Tradeoffs between transmission intervals, delays and performance. *IEEE Transactions on Automatic Control*, 55(8), 1781–1796.
24. Francis, B. A. (1987). Lecture notes in control and information sciences. *A Course in  $H_\infty$  Control Theory* (p. 88).
25. Xiao, L., Hassibi, A., & How, J. P. (2000). Control with random communication delays via a discrete-time jump system approach. In *Proceedings of the 2000 American Control Conference (ACC)* (Vol. 3, pp. 2199–2204). IEEE.

26. Soucek, S., Sauter, T., & Koller, G. (2003, November). Effect of delay jitter on quality of control in EIA-852-based networks. In *ECON'03. 29th Annual Conference of the IEEE Industrial Electronics Society* (Vol. 2, pp. 1431–1436). IEEE.
27. Xie, G., & Wang, L. (2004, December). Stabilization of networked control systems with time-varying network-induced delay. In 2004 43rd IEEE Conference on Decision and Control (CDC) (Vol. 4, pp. 3551–3556). IEEE.
28. Zhang, L., Shi, Y., Chen, T., & Huang, B. (2005). A new method for stabilization of networked control systems with random delays. *IEEE Transactions on Automatic Control*, 50(8), 1177–1181.
29. Liu, X. G., Martin, R. R., Wu, M., & Tang, M. L. (2006). Delay-dependent robust stabilisation of discrete-time systems with time-varying delay. *IEE Proceedings-Control Theory and Applications*, 153(6), 689–702.
30. Gao, H., Chen, T., & Lam, J. (2008). A new delay system approach to network-based control. *Automatica*, 44(1), 39–52.
31. Gao, H., & Chen, T. (2007). New results on stability of discrete-time systems with time-varying state delay. *IEEE Transactions on Automatic Control*, 52(2), 328–334.
32. Meng, X., Lam, J., Du, B., & Gao, H. (2010). A delay-partitioning approach to the stability analysis of discrete-time systems. *Automatica*, 46(3), 610–614.
33. Wang, J. (2015, August). A brief survey on networked control systems. In *2015 IEEE International Conference on Mechatronics and Automation (ICMA 2015)* (pp. 212–216). IEEE.
34. Baillieul, J., & Antsaklis, P. J. (2007). Control and communication challenges in networked real-time systems. *Proceedings of the IEEE*, 95(1), 9–28.

# Text-to-Image Encryption and Decryption Using Piece Wise Linear Chaotic Maps



K. Abhimanyu Kumar Patro, Shashwat Soni, V. K. Sharma and Bibhudendra Acharya

**Abstract** Generally, if an image is received as a Cipher, it is assumed that the data (to be sent) might be an image, but text can also be encrypted in the format of an image. So to confuse the attacker, this paper proposes a technique that encrypts the text into image using piecewise linear chaotic map (PWLCM). A text file can be taken as input and the proposed algorithm will convert it into a cipher image, which can then be converted back by the decryption process. The advantage of this scheme is that it is quite secure and hard to break. Also, PWLCM Maps is the simplest among chaotic maps. At first, the text is converted from 7-bit ASCII to its double equivalent and then padding is done to get the required matrix to form a structure for the image to which the text will be converted. After that permutation is done to the matrix bits and then diffusion occurs at two stages, first for the rows and then for the columns. Both permutation and diffusion are done using PWLCM map.

**Keywords** Security · Image encryption · Text to image · Piecewise linear chaotic map · SHA-256

## 1 Introduction

In the communication system, we exchange information in the form of text, images, audio, and videos. Cryptography is the process of making a piece of information inaccessible to all except the intended recipients of that information. In today's

---

K. A. K. Patro · S. Soni · V. K. Sharma · B. Acharya (✉)  
Department of Electronics and Telecommunication Engineering, National Institute of Technology Raipur, Raipur, India  
e-mail: [bacharya.etc@nitrr.ac.in](mailto:bacharya.etc@nitrr.ac.in)

K. A. K. Patro  
e-mail: [kpatro.phd2016.etc@nitrr.ac.in](mailto:kpatro.phd2016.etc@nitrr.ac.in)

S. Soni  
e-mail: [shashwatsoni@gmail.com](mailto:shashwatsoni@gmail.com)

V. K. Sharma  
e-mail: [vijay4247@gmail.com](mailto:vijay4247@gmail.com)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_4](https://doi.org/10.1007/978-981-15-0694-9_4)

world where electronic communication takes place via the Internet, particularly e-commerce, making sensitive data safe is quite important. Using different procedures of cryptography, data can be sent without the risk of the information being intercepted [1, 2]. Encryption is the basic means for covering large text information [3].

Images are most commonly used and therefore securing the images during transmission is important. To achieve this, we follow different conventional processes such as AES, DES, RSA, 3-DES, etc. [4, 5]. Unfortunately, many of them are cracked or no longer efficient in this modern world. They are not efficient to encrypt images of large size having high redundancy, huge data capacity and high correlation between adjacent pixels [6–13].

To overcome these problems, researchers have suggested different methods to achieve higher efficiency and stronger security to images and one of such process called ‘Chaos-based encryption technique’ was proposed by Matthews [14]. This technique helps in constructing secure cryptosystems. Since chaos maps are much sensitive to their initial conditions, they can be used to permute the image pixels in such a way that it is not easy to decrypt.

The chaotic maps used in this paper is piecewise linear chaotic maps (PWLCM). The proposed encryption scheme is a symmetric encryption system as the same key is used for encryption and decryption. In the present scheme, to prevent statistical and differential attacks, two different phases have been implemented and those are diffusion and permutation. Diffusion does not add extra data redundancies and at the same time provides better entropy. Permutation, on the other hand, intensifies the complexity between the pixels and key. In each algorithm, there are a defined number of stages for encryption and decryption. But many of them use one stage of encoding operation to change the order of the pixels of images which somehow gives chance for a cryptanalyst to break down the algorithm. This paper, however, shuffles the pixels of bit plane by permuting them followed by diffusion with two different keys, respectively, resulting in increased shuffling of pixels that results in increasing the level of security.

This paper has the following objectives:

- Two times encoding and decoding are done in this algorithm to shuffle the bits in bit-planes and to achieve higher keyspace.
- By using hash algorithm SHA-256, our algorithm is protected from Chosen-plaintext Attack (CPA) and Known-plaintext Attack (KPA).
- To increase the confusion and achieve a high diffusion rate among the pixels, PWLCM is utilized.

## 2 Preliminaries

### 2.1 Piecewise Linear Chaotic Map

PWLCM has been used to produce better results compared to other maps. It is mainly used in several encryption techniques. It is defined as

$$g_{n+1} = \begin{cases} \frac{g_n}{n} & \text{if } 0 \leq g_n < n \\ \frac{g_n - n}{0.5 - n} & \text{if } n \leq g_n < 0.5 \\ (1 - g_n) & \text{if } 0.5 \leq g_n < 1 \end{cases}$$

where  $g_n \in [0, 1]$  and  $n \in (0, 0.5)$  are the initial value and control parameter of PWLCM system, respectively.

#### Secure Hash Algorithm 256 (SHA-256)

Hash functions are used to provide integrity and authentication. In our scheme, we use SHA-256 and a 22261-length character read-only text file to generate a 128-bit key for generating initial key parameters for PWLCM. Even a slight change in the text file will completely change the hash value, thus this method of key generation provides high security against brute-force attack.

## 3 Proposed Methodology

Figure 2 illustrates the proposed methodology for the encryption of an image. A text file of length ' $L$ ' and padded so that its length is of the form ' $M*N$ ', where  $N$  is a multiple of 128 ( $2^7$  as a text is a 7-bit data). Then this data, which is in the form of a 1D matrix is then converted to a 2D matrix of rows ' $M$ ' and columns ' $N$ '. Now using this image, an SHA-256 key is generated using which 3 separate keys (explained in the next section) are created. Now using the first key, PWLCM Map is created and applied on Image which gives a permuted matrix on which PWLCM Map using key 2 is applied by XOR-ing it with the first column of the output of the previous step. Now the subsequent columns of the output are received by XOR-ing the original column with the output of the received 1D matrix (considered as the previous column of the output matrix). Now, this step is repeated with PWLCM Map using key 3 and the received 2D matrix is the Cipher Image.

Encryption steps are as discussed below:

### 3.1 Generation of Key Using SHA-256

Step-1: Using the SHA-256 hash algorithm, generate the 64-bit hexadecimal values. The 64-hex values are denoted as

$$\text{hash} = h_1, h_2, h_3, \dots, h_{63}, h_{64}$$

Step-2: Generate the keys of PWLCM system.

$$\text{msg} = \text{DataHash}(\text{IMAGE}, \text{SHA-256})$$

$$p1 = p - \left( \frac{\text{sum}(\text{msg}(1:10))}{10^{15}} \right) - \text{ceil} \left( \frac{\text{sum}(\text{msg}(1:10))}{10^{15}} \right) \times 0.01$$

$$z1(1) = z(1) - \left( \frac{\text{sum}(\text{msg}(11:20))}{10^{15}} \right) - \text{ceil} \left( \frac{\text{sum}(\text{msg}(11:20))}{10^{15}} \right) \times 0.01$$

$$p2 = p - \left( \frac{\text{sum}(\text{msg}(21:31))}{10^{15}} \right) - \text{ceil} \left( \frac{\text{sum}(\text{msg}(21:31))}{10^{15}} \right) \times 0.01$$

$$z2(1) = z(1) - \left( \frac{\text{sum}(\text{msg}(32:42))}{10^{15}} \right) - \text{ceil} \left( \frac{\text{sum}(\text{msg}(32:42))}{10^{15}} \right) \times 0.01$$

$$p3 = p - \left( \frac{\text{sum}(\text{msg}(43:53))}{10^{15}} \right) - \text{ceil} \left( \frac{\text{sum}(\text{msg}(43:53))}{10^{15}} \right) \times 0.01$$

$$z3(1) = z(1) - \left( \frac{\text{sum}(\text{msg}(53:64))}{10^{15}} \right) - \text{ceil} \left( \frac{\text{sum}(\text{msg}(53:64))}{10^{15}} \right) \times 0.01$$

where  $p1, p2, p3$  are generated system parameters,  $p$  is the original system parameter of PWLCM system, respectively,  $z1(1), z2(1), z3(1)$  are generated initial values and  $z(1)$  is the original initial value of PWLCM system.

### 3.2 Encryption Procedure

Step-1: Take a text I.

Step-2: Convert the 7-bit ASCII text into its double equivalent.

Step-3: Pad the converted double equivalent to resize the text into  $M*N$  values where M and N are integers. Let the matrix be denoted as 'IMAGE'.

Padding is done by the formula  $H(\text{end}+1:N*\text{ceil}(\text{numel}(H)/N)) = 0$ , where  $N = \text{floor}(L/128)$ , where 'L' is the length of the original text before padding and ' $M*N$ ' is the length of text after padding.

Step-4: Generate keys by iterating PWLCM map1 for  $M*N$  times, where  $M*N$  is the dimension of the padded matrix.

$$[\text{psort1}, \text{pindex1}] = \text{sort}(\text{map1}, \text{'descend'})$$

Step-5: Generate the permutation matrix by initializing a new matrix L1 of size  $1*(M*N)$  with zeros and then generate L1 matrix using PWLCM Map.



Step-6: Reshape L1 into an  $M*N$  image

Step-7: Using PWLCM map2, generate a sequence of length equal to the  $7*N$ .

$$[\text{psort2}, \text{pindex2}] = \text{sort}(\text{map2}, \text{'descend'})$$

Step-8: Set a threshold value (suppose 0.6)

Now if the value of a bit in the generated sequence of PWLCM map2 is above the threshold, set it as 1, otherwise 0.

Step-9: Convert the 7-bit binary newly generated map2, after thresholding, into the decimal matrix by taking 7 bits consecutively and converting them into corresponding decimal values.

Step-10: Create a new matrix L3 where the first column of the matrix is the bitwise-XOR of the columns of the L1 matrix, map2, and the next columns are bitwise-XOR of the previous column of L3 with the corresponding column of L1.

Step-11: Using PWLCM map3, we generate a sequence of length equal to the  $7*M$ .

$$[\text{psort3}, \text{pindex3}] = \text{sort}(\text{map3}, \text{'descend'})$$

Step-12: Set a threshold value (suppose 0.6)

Now if the value of a bit in the generated sequence of PWLCM map3 is above the threshold, set it as 1, otherwise 0.

Step-13: Convert the 7-bit binary newly generated map3 after thresholding into the decimal matrix by taking 7 bits consecutively and converting them into corresponding decimal values.

Step-14: Create a new matrix L4 where the first column of the matrix is the bitwise-XOR of the columns of L3 matrix and map3 and the next columns are bitwise-XOR of the previous column of L4 with the corresponding column of L3.

“L4 is the cipher image”.

## 4 Computer Simulations and Security Analysis

The proposed method is tested by doing analysis on a text file of length 22261 characters. The simulation process has been performed on a personal computer with i5 processor, 4 GB RAM, and MATLAB R2016a. Figure 1 illustrates the simulation results of the proposed scheme.

The security analyses are as follows:



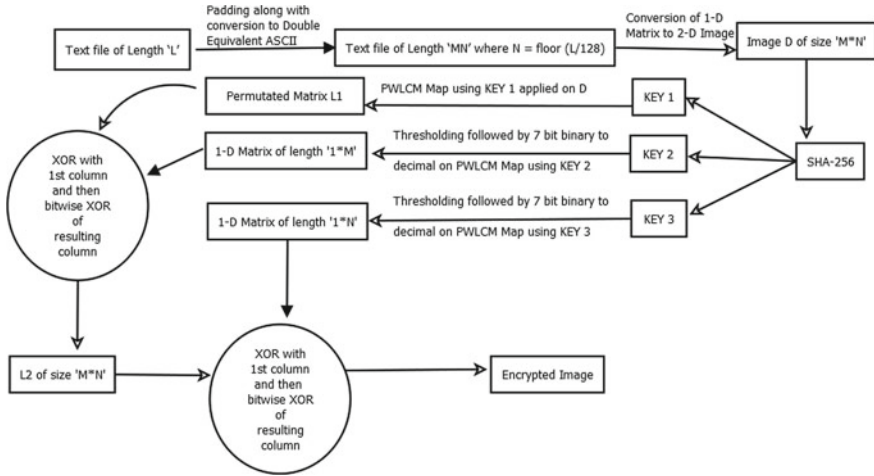


Fig. 2 Proposed cryptosystem block diagram

#### 4.1 Plaintext-Sensitivity Analysis

Our proposed cryptosystem is also sensitive to the plaintext. It is proved by modifying just one pixel in the plaintext, then generating the corresponding cipher image and finally generating the difference of original output encrypted image and the changed output encrypted image (Fig. 2).

#### 4.2 Key Space Analysis

The different keys that can be generated to use in this algorithm are as follows:

- Three different keys and initial values are used for each PWLCM system.
- SHA-256 hash algorithm is used of 256-bits.

Key space analysis is the situation where there are a large number of different keys that can be used, but out of them, only one key is appropriate. The chaotic map uses keys having a precision of  $10^{-15}$  [15]. The key space for SHA-256 hash algorithm is  $2^{128}$ .

So, from the above information, the total key space is  $(10^{15} \times 10^{15} \times 10^2) \times (10^{15} \times 10^{15} \times 10^{15} \times 10^{15}) \times (10^{15} \times 10^{15} \times 10^{15}) \times (10^{15} \times 10^{15}) \times 2^{128} = 1.6958 \times 2^{682}$  which is very larger than  $2^{128}$  to resist brute force attack [16].

The comparison results for key space are shown in Table 1.

**Table 1** Variations in key space results

Algorithm	Key space
Our algorithm	$1.6958 \times 2^{682}$
Ref. [1]	$2^{128}$

### A. Statistical Attacks Interpretation

Frequency Analysis: This method is based on histogram analysis. The histogram shows the number of times a letter appears in a text. If the histogram of the ciphertext has all letters in a regular way, the algorithm could prevent this kind of attack. On the other hand, in case of irregular histogram curve, the attacker may get a clue to find the message by doing some kind of frequency attack. That means, the message becomes vulnerable even after encryption. Figure 3 illustrates the histogram output of a text file of the length of 22261 characters. From the output, we can say that there is a regular distribution of gray levels in encrypted image. This proves that this algorithm has stronger resistant towards the statistical attacks. The lesser is the variance, the better is the proposed scheme.

### B. Differential Attack Analysis

The most important thing needed with the encryption techniques is that the cipher image should be quite different from the original image. To quantify the difference between an encrypted image and original text file, two measures were used: the Number of Pixel Changing Rate (NPCR) and Unified Average Changing Intensity (UACI) [6, 7]. These are the two most important methods using which evaluation of the strength of image encryption is done for differential attacks. Table 2, illustrates the NPCR and UACI values for 3 sets of text files and their corresponding Cipher Images  $C^1$ ,  $C^2$ , and  $C^3$ .

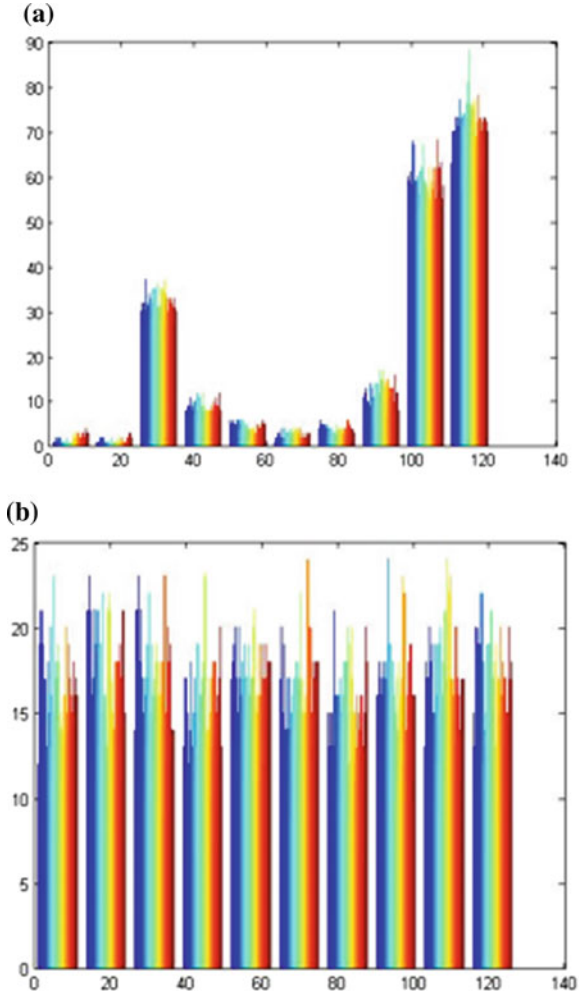
### C. Information Entropy Analysis

The more is the value of information entropy, more will be the pixels randomness that can be achieved. When the plaintext is sent in the diffusion stage, the value of all the symbols must be modified. The entropy of the message should be  $H(m) = N$ , if a message is ciphered with  $2N$  possible values. In our method, the maximum obtained entropy is 6.56. The input text entropy (from Fig. 3a) 0.02. The ciphertext entropy (from Fig. 3b) is 6.49. This means that therefore, the diffusion process is strong for doing any kind of analysis.

### D. Known-plaintext Attack (KPA) and Chosen-plaintext Attack (CPA) Analysis

The idea of such dangers is that, if the hacker has access to either a part of plaintext/ciphertext or its corresponding ciphertext/plaintext respectively then he/she will know how to decrypt the text. Since we know that the keys that are generated in this algorithm are brought out from the plaintext taking use of hash values, so these are temporary keys for every unique input of text. Even if the hacker has some chosen plain/ciphertext, it is not possible to generate all other keys (temporary) to decipher the whole text. Due to its high dependability over the input text, which changes every time makes it more secure toward CPA and KPA.

**Fig. 3** Histogram outputs of: **a** original text file  
**b** cipher image of text



**Table 2** NPCR and UACI values

Input image	NPCR	UACI
Cipher image $C^1$ and $C^2$	99.5926	33.4740
Cipher image $C^2$ and $C^3$	99.6124	33.5410

## 5 Conclusion

As indicated by security analysis the proposed encryption scheme has a large key space. The frequency and histogram analysis parameters are also within the desirable range. Differential analysis parameters also lie within the desirable range. These factors indicate that our encryption is highly secure and can be applied for gray images of any size. The decryption process is very simple, which is done in the reverse procedure of encryption by using the accurate keys. Change in even a single bit will lead to error. The proposed method is easy to implement and stands strong against the cyber attacks.

## References

1. Murillo-Escobar, M. A., Abundiz-Pérez, F., Cruz-Hernández, C., & López-Gutiérrez, R. M. (2014). A novel symmetric text encryption algorithm based on logistic map. In *2014 International Conference on Communications, Signal Processing and Computers* (Vol. 4953).
2. Singh, S., & Jain, A. (May 2013). An enhanced text to image encryption technique using RGB substitution and AES. *International Journal of Engineering Trends and Technology (IJETT)*, 4(5).
3. Abusukhon, A. (May–June 2013). Block cipher encryption for text to image algorithm. *International Journal of Computer Engineering and Technology (IJ CET)*, 4(3).
4. Coppersmith, D. (1994). The data encryption standard (DES) and its strengths against attacks. *IBM Journal of Research and Development*, 38(3), 243–250.
5. Pub NF. 197: Advanced encryption standard (AES). (2001). Federal information processing standards publication, 197, 311–441.
6. Patro, K., & Acharya, B. (2018). Secure multi-level permutation operation based multiple colour image encryption. *Journal of Information Security and Applications*, 40, 111–133.
7. Patro, K., Banerjee, A., Acharya, B. (2017). A simple, secure and time efficient multi-way rotational permutation and diffusion based image encryption by using multiple 1-D chaotic maps. In *International Conference on Next Generation Computing Technologies* (pp. 396–418). Singapore: Springer.
8. Gao, H., Zhang, Y., Liang, S., & Li, D. (2006). A new chaotic algorithm for image encryption. *Chaos, Solitons and Fractals*, 29, 393–399.
9. Samhita, P., Prasad, P., Patro, K. A. K., & Acharya, B. (2016). A secure chaos-based image encryption and decryption using crossover and mutation operator. *International Journal of Control Theory and Applications*, 9(34), 17–28.
10. Gupta, A., Thawait, R., Patro, K. A. K., & Acharya, B. (2016). A novel image encryption based on bit-shuffled improved tent map. *International Journal of Control Theory and Applications*, 9(34), 1–16.
11. Shadangi, V., Choudhary, S. K., Patro, K. A. K., & Acharya, B. (2017). Novel Arnold scrambling based CBC-AES image encryption. *International Journal of Control Theory and Applications*, 10(15), 93–105.
12. Brindha, M., & Ammasai Gounden, N. (2016). A chaos-based image encryption and lossless compression algorithm using hash table and Chinese remainder theorem. *Applied Soft Computing*, 40, 379–390.
13. Wu, X., Kurths, J., & Kan, H. (2017). A robust and lossless DNA encryption scheme for color images. *Multimedia Tools and Applications*.
14. Matthews, R. (1989). On the derivation of a chaotic encryption algorithm. *Cryptologia*, 13(1), 29–42.

15. Floating-Point Working Group. (1985). IEEE Standard for Binary Floating-Point Arithmetic. ANSI. IEEE Std (pp. 754–1985).
16. Kulsoom, A., Xiao, D., & Abbas, S. A. (2016). An efficient and noise resistive selective image encryption scheme for gray images based on chaotic maps and DNA complementary rules. *Multimedia Tools and Applications*, 75(1), 1–23.

# Security Threats, Attacks, and Possible Countermeasures in Internet of Things



Shams Tabrez Siddiqui, Shadab Alam, Riaz Ahmad and Mohammed Shuaib

**Abstract** The idea to connect everything to anything and at any point of time is what vaguely defines the concept of Internet of Things (IoT). The concept of IoT is not only about providing connectivity but also facilitating interaction among these connected things. Though the term IoT was introduced in 1999 but has drawn significant attention during the past few years. The pace at which new devices are being integrated into the system will profoundly impact the world in a good way but also poses some serious threats with regard to security and privacy. IoT in its current form is susceptible to a multitudinous set of attacks. One of the greatest concerns of IoT is to provide security assurance for the data exchange because data is vulnerable to a number of attacks by the attackers at each layer of IoT. The IoT has layered structure, where each layer provides a service. The security vary from layer to layer as each layer serves a different purpose. The aim of this paper is to analyze the various security and privacy threats related to IoT. Furthermore, this paper also discusses numerous existing security protocols operating at different layers, potential attacks, and suggested countermeasures.

**Keywords** Internet of things · Attacks · Security · Threats · Protocols

---

S. T. Siddiqui · S. Alam · M. Shuaib  
Department of Computer Science, Jazan University, Jazan, Saudi Arabia  
e-mail: [stabrezsiddiqui@gmail.com](mailto:stabrezsiddiqui@gmail.com)

S. Alam  
e-mail: [s4shadab@gmail.com](mailto:s4shadab@gmail.com)

M. Shuaib  
e-mail: [talkshuaib@gmail.com](mailto:talkshuaib@gmail.com)

R. Ahmad (✉)  
Department of Computer Science, Aligarh Muslim University, Aligarh, India  
e-mail: [riaz.ahmad.tech@gmail.com](mailto:riaz.ahmad.tech@gmail.com)



## 1 Introduction

IoT emerged in the year 1999 with the introduction of Wireless Sensor Networks (WSN) and technologies like Radio-Frequency Identification (RFID). The concept behind the IoT is to connect everything to anything, anywhere, and at any moment of time. For making physical or virtual connections, it uses objects like sensors, actuators, etc. The success of IoT infrastructure and applications depends on IoT security. IoT collects the data from a vast geographical region using sensors and actuators [1].

The IoT is going to gain the attention of masses. The concept of IoT devices is not only about providing connectivity but also they need to be interactive. The need of hour is that they should deploy context-based interactions [2]. There will be billions of interconnectivity among the internet that will surely open doors for hackers and with that there will be a lot of security and privacy threats that will need immediate supervisions.

The objective of IoT technology is to provide interconnections between humans, things, and between humans and objects. In the IoT infrastructure, the sensors and objects are integrated for communications that can work successfully without human interventions. The sensors play an important role in IoT as these devices not only collect heterogeneous data but also monitors the data with diversity and is quite intelligent and dynamic in nature [3, 4]. The major IoT principles include confidentiality, authentication, availability, heterogeneity, lightweight solutions, key management, policies, and integrity.

IoT has a layered structure where each layer provides a service. Usually, the IoT architecture is categorized in three layers, namely, application, network, and perception layer. The security issues like privacy, authorization, verification, access control, system configuration, information storage, and management that are the real challenges of the IoT infrastructure [5, 6]. The security needs vary from layer to layer as each layer serves a different purpose [5]. Undoubtedly, to make IoT a reality the security issues need to be resolved. There are two types of security challenges, namely, technological and security challenges. The technological challenges include wireless technologies and the distributed nature of the IoT. The challenges related to authentication and confidentiality included in the security [7].

This paper discusses the protocols present on different IoT layers and identify the security threats at each layer. Different security issues and its countermeasures have been discussed in detail. The objective of this paper is to enlighten the essential security protocols of IoT that obliging for the prevention of harmful threats.

## 2 IoT Architecture

IoT has a three-layered architecture. The three layers are as follows:

**Table 1** Different protocols that are present on different layers

IoT layers	Protocols
Application layer	CoAP, DDS, MQTT, SMQTT, AMQP
Network layer	6LoWPAN, RPL, CORPL, CARP, 6TISCH
Perception layer	LTE-A, Z-Wave, ZigBee smart, DASH7, 802.11AH

- The Application Layer,
- The Network Layer, and
- The Perception Layer.

*The Application Layer:* The main aim of the application layer is to deliver specific services to its users [8]. It defines numerous applications of IoT, viz., smart home, health, cities where it can be deployed.

*The Network Layer:* This layer is most prone to attacks, it aggregates data from existing infrastructures and transmits the data to other layers. It processes the sensor data. The major security issues usually related to authentication and integrity of data that is being transmitted [9].

*The Perception Layer:* This is the physical layer, even known as the lowest layer of the IoT architecture and reflected as a brain of the three-layered architecture. The sensing devices like the sensors and actuators are present at this layer. This layer is also known as the sensor layer [10, 11] (Tables 1 and 2).

### 3 Security Requirements

IoT infrastructure consists of a lot of personal information such as name, date of birth, locations, etc. Therefore, we need to provide strict measures to protect the data and tackle privacy risks. In order to overcome the security challenges, the layered structure is adopted. The basic security properties that need to be implemented are confidentiality, authenticity, integrity, and availability. There are a number of other security requirements that are derived from the basic security requirements such as scalable, IP Protocol-Based IoT, Heterogeneous IoT, and Lightweight Security.

### 4 IoT Security Threats

The threats can broadly be classified into three categories. The categories are capture, disrupt, and manipulate. The capture threat means capturing information or system without authorization. The capture threats are such threats that are designed to gain access of information that is either logical or physical on a system. The disrupt

**Table 2** Application, network, and perception layer protocols

PROTOCOLS	PURPOSE
CoAP	Constrained application protocol (CoAP) is designed in such a way that it enables the low-power sensors to make usage of restful services. It is very much similar to HTTP and is built upon the UDP instead of TCP packets [12]
DDS	Data distribution service (DDS) provides an excellent quality of service that can have scalability with excessive overall performance and reliability that suits the IoT and M2M communication [12]
MQTT	Message Queue Telemetry Transport Protocol (MQTT) facilitates the embedded connectivity between applications and the middlewares at one side whereas the networks and communications on the other side [13]
SMQTT	Secure Message Queue Telemetry Transport Protocol (SMQTT), the message is encrypted before delivering to multiple nodes in the network [14]
AMQP	Advanced Message Queuing Protocol is a software layer protocol having three additives, namely, exchange, message queue, and binding. This protocol is generally message-oriented for middleware environment [15]
6LoWPAN	Wireless sensor network is one of the applications of IPv6 Low-Power Wireless Personal Area Network (6LoWPAN) system, uses it while sending data as a packet. It provides huge variety of network connected to internet providing end-to-end services [16]. The specification supports different length addresses, low bandwidth, different topologies including star or mesh, power consumption, low cost, scalable networks, mobility, unreliability, and long sleep time
RPL	Routing Protocol for wireless network with Low-Power consumption having Lossy Networks (RPL) supports one-to-one communication [16]. It can quickly create network routes, adapt topology in an efficient way, share routing knowledge but susceptible to packet loss
CORPL	It is a routing protocol for cognitive radio enabled AMI network? An extension of RPL designed for the cognitive networks but with two new modification that uses DODAG topology generation [17]
CARP	Channel-Aware Routing Protocol (CARP) is a distributed routing protocol designed for light-weighted packets in IoT. Therefore, it is used for acoustic communication under the water [18]
6TiSCH	IPv6 time-slotted channel hopping (6TiSCH) working group in IETF is developing standards to allow IPv6 to pass through TSCH mode of IEEE 802.15.4e data links [19]. TSCH demonstrate end-to-end reliability. This essentially a MAC layer that offers globally synchronized mesh network of sleepy node and is also defined as minimal configuration
LTE-A	Long-Term evolution advanced (LTE-A) is an agglomeration of cellular network. As compared to other cellular networks it is one of the most scalable and lower cost protocol [20]

(continued)

**Table 2** (continued)

PROTOCOLS	PURPOSE
Z-WAVE	Z-Wave is a low cost and low-power MAC protocol that design aimed specifically for home automation [21]
ZigBee Smart Energy	An enhancement to the customary ZigBee is ZigBee IP or Smart which is designed for the substantial range of IoT applications including smart homes, healthcare systems, and for remote controls. It supports numerous topologies including star, peer-to-peer or cluster tree [22]
DASH7	DASH7 is a wireless communication protocol for active RFID specifically designed for scalable, long-range outdoor coverage with higher data rate. It provides low cost and light-weighted solutions [23]
IEEE 802.11 AH	IEEE 802.11ah is a wireless networking protocol with low energy capable communication standard [24]

**Table 3** The description of threats at each layer

IoT layers	Threats
Application layer	Malicious code attacks, Tampering with node-based applications, Inability to receive security patches, Hacking into the smart meter/grid, Phishing Attack, Malicious Virus/worm, Malicious Scripts, Remote configuration, Mis-configuration, Security management, Management system
Network layer	DoS attack, Gateway attacks, Unauthorized access, Storage attacks, Injecting fake information, Spoofing attacks, Sinkhole attacks, Wormhole attacks, Man-in-the-Middle attack, Routing attacks, Sybil attacks, Unauthorized access
Perception layer	Wireless Sensor Networks (WSN), Eavesdropping, Repudiation, Noise in data, Privacy threats services abuse, RFID, Service information Manipulation, Sniffing attacks, Identity masquerade, Replay attack

threat means denying access or destroying a system. The manipulated threat means manipulating time series data, identity, or the data (Table 3).

## 5 IoT Challenges

Due to the vast scale of IoT infrastructure with a huge number of devices involved in developing a successful IoT application is not an easy task and have to face a lot of challenges. Some of the challenges are, namely, mobility, reliability, availability Identification, scalability, data integrity, management, energy management, interoperability, and security and privacy.

*Mobility:* It is one of the essential issues of the IoT paradigm. As IoT devices move freely from one network to another, therefore, movement detection is important to monitor the device location and respond to the topology that changes accordingly due to which layer of complexity escalate to another level [25].

*Reliability:* Reliability is a very critical requirement in the application that requires all the emergency responses correctly otherwise, it will be a huge disastrous scenario. In IoT applications, data collection, communication should be fast and highly reliable [25].

*Scalability:* Other challenges of IoT application is scalability, where enormous number of devices are connected to a network, therefore, the protocols must have efficient extensible services to meet the IoT devices requirements [26].

*Management:* Managing a vast number of devices and keeping track of their failures, configurations, and performances in the network is an immense challenge [26].

*Energy management:* In IoT devices, energy is required still not adequately met. Some routing protocols at an early stage of development supports low power communication but to make IoT devices more power efficient, Green technology must be employed [25].

*Availability:* Availability means the service subscriber provides the service anytime and anywhere for the service subscribers. Software service provided to anyone who is authorized to, whereas the hardware availability means easy to access and are compatible with IoT functionality and protocols.

*Interoperability:* Huge number of heterogeneous devices and protocols work with each other. This becomes a challenging task due to the number of IoT devices using various platforms [25].

*Identification:* To provide innovative services, the IoT devices are interconnected with numerous objects, and hence, an efficient naming and identity managing system is required to specify the object [26].

*Data Integrity:* IoT devices are heterogeneous in nature, therefore, they have to deal with big amount of data. Handling big data is very crucial as overall the performance is directly proportional to the features of data management services. Became more complicated when data integrity features are considered, it also affects the QoS, Privacy, and Security related issues specifically on outsourced data [25, 26].

## 6 Counter Measures

The countermeasures that can be taken are the authentication measures, establishment of trust, and acceptance of federated architecture awareness of security issues (Table 4).

**Table 4** The countermeasure of threats at each layer

IoT Layers	Protocols	Threats	Countermeasures	Countermeasures description
Application layer	CoAP, DDS, MQTT, SMQTT, AMQP	Malicious code attacks	Runtime type checking, Firewall checks	Seem to do runtime type checking, immune for all ill-typed code tried. At runtime, the firewall checks have to be done
		Tampering with node-based applications	Physically secure design	Physically secure designing of devices should not be of high quality and unreliable [27]
		Inability to receive security patches	Evading security risks with regular patching and support services	
		Hacking into the smart meter and kill the grid	Security Frameworks to Prevent from Hacking the Grid	
		Malicious injection	Custom FileZilla as the FTP client	The credentials of the websites stored in plain text by FileZilla
		Remote configuration	Configuring and managing VPNs	NCP engineering offers inclusive software that designed for the clients indispensable to control large networks
		Application security	Web Application Scanner	Discovery of various threats which is present on the front end of web [28]
		Security management	Security management is the identification of an organization's assets followed by the development, documentation, and implementation of policies and procedures for protecting these assets	
		Data security	Fragmentation redundancy scattering	Data on cloud splits and apportions to various fragments for the storage in servers [29]
		Shared resources	Holomorphic encryption	Ciphertext allowed to reckon immediately without decryption [27]
Mis-configuration	This attack can occur at any level of an application stack including the platform, application server, web server, database, and framework [30]			

(continued)

**Table 4** (continued)

IoT Layers	Protocols	Threats	Countermeasures	Countermeasures description
Network layer	6LoWPAN, RPL, CORPL, CARP, 6TISCH	DoS attack	This can be handled by assuring that resources are committed to a client only after proper authentication, utilization of proxy servers with sufficient resources, protocol scrubbing (to remove protocol uncertainties which can be misused for attacks)	
		Gateway attacks	Blocking spyware at the Network gateway	Block against viruses, spam, and intruders, organizations deploy countermeasures at the network gateway and again in individual client systems
	Unauthorized access	Device authentication	Without any authentication, the device cannot enter or connect with other nodes in the IoT system	
	Storage attacks	In case of physical security weaknesses, the attackers can effortlessly access the storage medium via disassemble the device		
	Injecting fake information	Injecting fake routing control packets in the network		
	Spoofing attacks	IPsec will significantly cut down on the risk of spoofing	Use authentication based on the key exchange between the machines on your network; Enable encryption sessions on your router so that trusted hosts that are outside your network can securely communicate with your local hosts	
	Sinkhole attacks	Security aware and ad hoc routing	Stops inside attacks from the network of IoT, use key management, authentication, and geographical routing protocols, and drop adversary from the network	
	Wormhole attacks	Routing Protocol (AODV and DSR)	Stratagem the packet LEACH techniques for detecting and thus defending against said attacks	

(continued)

**Table 4** (continued)

IoT Layers	Protocols	Threats	Countermeasures	Countermeasures description
Perception Layer	LTE-A, Z-Wave, Zigbee smart, DASH7, 802.11ah	Man in the Middle attack	Secure/Multipurpose Internet Mail Extensions, or S/MIME; Authentication Certificates	Hackers will never go away, but one thing you can do is make it virtually impossible to penetrate your systems by implementing Certificate-Based Authentication for all employee machines and devices
		Routing information attacks	Encrypting routing tables	was identifies different security issues on the web by encryption process in rout
		Sybil attacks	Authentication and encryption preclude from outsider attack, Privilege Attenuation, Economic Incentives, public-key cryptography preclude from insider attacks [25, 31]	
		Unauthorized access	Two-factor authentication, IP White listing	
		RF interface on RFID	Device authentication	Before sending and receiving of data from a new physical device the device should authenticate itself
		Jamming node in Wireless Sensor Networks (WSN)	IPsec Security channel	Can be circumvented by stratagem different paths for routing [25, 31]
		Eavesdropping	Session Keys protect NPDU from Eavesdropper [31]	
		Sniffing attacks	Sniffer detection tools like ARP Watch, PromiScan, Anti-Sniff, Pro detect	Applications using secure protocols viz., HTTPS, SFTP, SSH. If obligatory than VPN can be used to provide the users with secure access
		Noise in data		
		Privacy threats	RFID	

(continued)



Table 4 (continued)

IoT Layers	Protocols	Threats	Countermeasures	Countermeasures description
		Services abuse	verify identity; strong password	Generally, a unique user ID is assigned to each user, but passwords are something you must set (or change) by yourself. If your User ID and Password are compromised or stolen, somebody else might use them to access your system or other systems, masquerading as a legitimate user
		Identity masquerade		
		Service information manipulation		
		Reputation	Create secure audit trails; Use digital signatures	
		Replay attack	Timestamps, one-time passwords, and challenge-response cryptography [25]	

## 7 Conclusion

IoT has recently emerged as an important research topic. Due to emerging technology attackers take advantages of the IoTs great potential to threaten users privacy, security, and wide variety of attacks. Therefore, it is essential to focus on the security parameters and heeded toward giving new feasible solutions to block all possible threats and vulnerabilities to IoT. This paper presents a comprehensive overview of security threats and attacks on IoT. Application, network and perception layer protocols with purpose been discussed. In addition, this paper suggested several countermeasures against identified security threats of each layer.

A lot more need to happen in near future in the area of IoT applications. This IoT field will definitely mature the impact of human life in inconceivable ways over the next decades. As IoT is going to play an indispensable part in our lives, steps should be taken to ensure the security and privacy of the users.

Future work involves finding alternative solutions for attacks that are less complex and less time-consuming. Future research involves development of protocols and finds ways to overcome security threats and attacks.

## References

1. Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, 29(7), 1645–1660.
2. Roman, R., Najera, P., & Lopez, J. (2011). Securing the internet of things. *Computer*, 9, 51–58.
3. Horrow, S., & Sardana, A. (2012). Identity management framework for cloud based internet of things. In *Proceedings of the First International Conference on Security of Internet of Things* (pp. 200–203). ACM.
4. Whitmore, A., Agarwal, A., & Da Xu, L. (2015). The Internet of Things—A survey of topics and trends. *Information Systems Frontiers*, 17(2), 261–274.
5. Aazam, M., St-Hilaire, M., Lung, C. H., & Lambadaris, I. (2016). PRE-Fog: IoT trace based probabilistic resource estimation at Fog. In *2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC)* (pp. 12–17). IEEE.
6. Jiang, H., Shen, F., Chen, S., Li, K. C., & Jeong, Y. S. (2015). A secure and scalable storage system for aggregate data in IoT. *Future Generation Computer Systems*, 49, 133–141.
7. Li, S., Tryfonas, T., & Li, H. (2016). The Internet of Things: A security point of view. *Internet Research*, 26(2), 337–359.
8. Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., & Ayyash, M. (2015). Internet of things: A survey on enabling technologies, protocols, and applications. *IEEE Communications Surveys & Tutorials*, 17(4), 2347–2376.
9. Pongle, P., & Chavan, G. (2015). A survey: Attacks on RPL and 6LoWPAN in IoT. In *2015 International Conference on Pervasive Computing (ICPC)* (pp. 1–6). IEEE.
10. Tsai, C. W., Lai, C. F., & Vasilakos, A. V. (2014). Future Internet of Things: Open issues and challenges. *Wireless Networks*, 20(8), 2201–2217.
11. Sethi, P., & Sarangi, S. R. (2017). Internet of things: Architectures, protocols, and applications. *Journal of Electrical and Computer Engineering*.
12. Karagiannis, V., Chatzimisios, P., Vazquez-Gallego, F., & Alonso-Zarate, J. (2015). A survey on application layer protocols for the internet of things. *Transaction on IoT and Cloud Computing*, 3(1), 11–17.

13. Locke, D. (2010). Mq telemetry transport (mqtt) v3. 1 protocol specification. *IBM developer Works Technical Library*.
14. Singh, M., Rajan, M. A., Shivraj, V. L., & Balamuralidhar, P. (2015). Secure mqtt for internet of things (iot). In *2015 Fifth International Conference on Communication Systems and Network Technologies* (pp. 746–751). IEEE.
15. OASIS, O. S. (2012). OASIS Advanced Message Queuing Protocol (AMQP) Version 1.0. Burlington, MA, USA: OASIS.
16. Winter, T., Thubert, P., Brandt, A., Hui, J., Kelsey, R., Levis, P., & Alexander, R. (2012). RPL: IPv6 routing protocol for low-power and lossy networks (No. RFC 6550).
17. Aijaz, A., & Aghvami, A. H. (2015). Cognitive machine-to-machine communications for Internet-of-Things: A protocol stack perspective. *IEEE Internet of Things Journal*, 2(2), 103–112.
18. Zhou, Z., Yao, B., Xing, R., Shu, L., & Bu, S. (2016). E-CARP: An energy efficient routing protocol for UWSNs in the internet of underwater things. *IEEE Sensors Journal*, 16(11), 4072–4082.
19. Dujovne, D., Watteyne, T., Vilajosana, X., & Thubert, P. (2014). 6TiSCH: Deterministic IP-enabled industrial internet (of things). *IEEE Communications Magazine*, 52(12), 36–41.
20. Hasan, M., Hossain, E., & Niyato, D. (2013). Random access for machine-to-machine communication in LTE-advanced networks: Issues and approaches. *IEEE Communications Magazine*, 51(6), 86–93.
21. Yassein, M. B., Mardini, W., & Khalil, A. (2016). Smart homes automation using Z-wave protocol. In *2016 International Conference on Engineering & MIS (ICEMIS)* (pp. 1–6).
22. Wang, C., Jiang, T., & Zhang, Q. (2016). *ZigBee® network protocols and applications*. Auerbach Publications. 604 pp.
23. Cetinkaya, O., & Akan, O. B. (2015). A DASH7-based power metering system. In *2015 12th Annual IEEE Consumer Communications and Networking Conference (CCNC)* (pp. 406–411). IEEE.
24. <https://standards.ieee.org/standard/802.11ah-2016.html>.
25. Salman, T., & Jain, R. (2017). *Networking Protocols and Standards for Internet of Things*. Wiley.
26. Triantafyllou, A., Sarigiannidis, P., & Lagkas, T. D. (2018). Network protocols, schemes, and mechanisms for internet of things (iot): Features, open challenges, and trends. *Wireless Communications and Mobile Computing*.
27. Abomhara, M., & Kjøien, G. M. (2014). Security and privacy in the Internet of Things: Current status and open issues. In *2014 International Conference On Privacy And Security In Mobile Systems (Prisms)* (pp. 1–8). IEEE.
28. Zhang, Z. K., Cho, M. C. Y., Wang, C. W., Hsu, C. W., Chen, C. K., & Shieh, S. (2014). IoT security: Ongoing challenges and research opportunities. In *2014 IEEE 7th International Conference On Service-Oriented Computing And Applications* (pp. 230–234). IEEE.
29. Migault, D., Palomares, D., Herbert, E., You, W., Ganne, G., Arfaoui, G., & Laurent, M. (2012). E2e: An optimized ipsec architecture for secure and fast offload. In *2012 Seventh International Conference on Availability, Reliability and Security* (pp. 365–374). IEEE.
30. <https://support.portswigger.net/customer/portal/articles/1965728-using-burp-to-test-for-security-misconfiguration-issues>.
31. El Mouaatamid, O., Lahmer, M., & Belkasmi, M. (2016). Internet of Things Security: Layered classification of attacks and possible Countermeasures. *Electronic Journal of Information Technology*, (9).

# Securing IoT-Driven Remote Healthcare Data Through Blockchain



Sarthak Gupta, Virain Malhotra and Shailendra Narayan Singh

**Abstract** Blockchain is the latest technology which is used in cryptocurrencies such as bitcoin and ether. Blockchain is the decentralized distributed ledger, which is based on the peer-to-peer network method. As the blockchain is mainly developed for implementation in the virtual cryptocurrency such as bitcoin, so its main purpose is clear that is security. Even though the healthcare industry is leading in majority of fields whether it is technology, equipment, researches, medicines, etc. We have even reached to remote locations through IoT devices and but one major thing is still lacking that is security of the data. Huge amount of data is being generated everyday from patient's medical checkups, treatments, symptoms, etc., which is to be dealt with care as they are very crucial and can be tempered by hackers which can lead to serious problems. Therefore, such critical data must need to be secured with blockchain as it makes it very difficult to tamper with data. This paper deals with the implementation of the blockchain to solve the above-mentioned problems.

**Keywords** Ethereum · Smart contracts · Blockchain · IoT · IPFS

## 1 Introduction

Presently, there is a huge advancement in technology. There is advancement in various fields such as agriculture, space, automobiles, etc. But the most important advancement has been made in the field of health care. The new technologies such as IoT have helped in covering and monitoring the health of remote population. People who don't have access to doctors due to lack of availability of doctors in such areas can

---

S. Gupta · V. Malhotra · S. N. Singh (✉)  
ASET, Amity University, Noida, India  
e-mail: [snsingh36@amity.edu](mailto:snsingh36@amity.edu)

S. Gupta  
e-mail: [guptasarthak03@gmail.com](mailto:guptasarthak03@gmail.com)

V. Malhotra  
e-mail: [virain.malhotra3@gmail.com](mailto:virain.malhotra3@gmail.com)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_6](https://doi.org/10.1007/978-981-15-0694-9_6)

get their health checkups on regular basis with the help of IoT and also get recommendations and prescription based on the data retrieved during the checkups. All this is possible while the patients don't have any means to contact the doctor face to face. All the data respective to the patient collected during the checkups or during the treatment which includes BP level, pulse rate, ECG, etc. are stored on the website. Both patient and doctors can access those data to monitor and analyze the patient medical history and recovery.

But with the advancement in the technology and huge amount of data is being generated and there comes the threat of tampering of data which can be fatal especially in the case of health care. As the data can be misused this can harm the patients' health. Hackers can also modify the data which will result in the distortion of the treatment going on. Since the security of the data is very essential in the healthcare applications, therefore, the data must be dealt with care and must be prevented from any kind of data tampering with the most secured technology such as blockchain.

Blockchain is the decentralized distributed ledger, which is based on the peer-to-peer network method. As the blockchain is mainly developed for implementation in the virtual cryptocurrency such as bitcoin, so its main purpose is clear that is security. To hack or corrupt the data secured by the blockchain, one need to change entire chain of blocks, which requires huge computational power and therefore is very difficult to do. This paper deals with the implementation of the blockchain in securing the healthcare data and preventing it from data manipulations by hackers.

## 2 Literature Survey

Bhabendu Kumar Mohanta states in his paper about the various components and the principles on which smart contract works. Moreover, he also states the various application and uses cases of smart contracts in blockchain [1].

You Sun has proposed a decentralized attribute-based scheme in his paper for blockchain implication in healthcare system for effective verification of authenticity of signer's identity. Even holistic on-chain and off-chain collaborative storage system was proposed in his paper [2].

Tran Le Nguyen, in his paper, proposed an application to store and create a database for doctors and patients and this paper is based on simulation space conceptual model. In this paper, he proposed bitcoin to be paid as a payment to the doctors [3].

Culver [4] indicated that collecting information to audit Medical Loss Ratio (MLR) and improper payment for subsidized programs creates substantial overhead for both providers and health care plans. This paper suggests an architectural model which is a solution to groups like providers, health plans and government which are the main stakeholders for payment process. The benefit of this model is the majority of health plan process will enable the providers to submit claims and providers to

submit claims and provide other stakeholders through 15 nodes in blockchain. Each node has smart contract and the same so that all stakeholders can view applicable data or interface directly with the Blockchain to executive the agreed-upon contracts. One more advantage of this model is the third-party nodes are present in nodes outside of the three main stakeholders [4].

Atlam's paper provides an overview of the integration of the blockchain with the IoT by highlighting the integration benefits and challenges. The future research directions of blockchain with IoT are also discussed. In his paper, he concluded that the combination of blockchain and IoT may provide a powerful approach which can significantly pave the way for new business models and distributed applications [5].

### **3 Proposed Model**

Looking at the problem of insecurity of data in the above existing system, the proposed system introduces the concept of blockchain. Blockchain is the decentralized distributed ledger, which is based on the peer-to-peer network method. It is a global online database which anyone, anywhere with an internet connection can use. Unlike traditional databases which are maintained and third party based, the blockchain doesn't belong to anyone. Blockchain stores information permanently across a network of personal computers, this not only decentralize the network but distributes it too [6]. Every new block which is added in the blockchain is shared with the other blocks with the timestamp and thus each block in the blockchain contains the information of the other blocks. This makes the blockchain hack-proof and difficult to tamper with.

#### ***3.1 Integration of IoT with Blockchain***

IoT is making huge advancement in wireless communication, sensor-based technology and if we combine it with the technologies like big data and Artificial Intelligence it makes the system more intelligent while not exceeding the cost. But taking into consideration the limited maintenance cost and management cost, there is restricted privacy of data and also insecure exchange of data among the personal computers. There comes concept of Blockchain into play [7]. Blockchain which is based on the distributed ledger technology can be implemented to IoT networks which themselves are distributed in nature. Therefore, these networks can be secured and shielded from any kind of data tampering at any point [8].

### 3.2 Blockchain in Health Care

Remote healthcare monitoring and analysis requires cloud storage for resilience and easy access of the data retrieved. Even though the cloud is the best platform for privacy and sharing of data among various subjects involved in healthcare monitoring analysis such as patients, doctors, data analyst, etc., it does not support interoperability among the above-mentioned stakeholders and also it does not guarantee the integrity and authenticity of medical data [9]. So, to mitigate the above flaws, blockchain technology can be incorporated in this model which ensures and enhance integrity, consistency, and also authenticity of the medical records stored [10]. High security and confidentiality of medical data is the first and foremost thing of concern for the patients and all this data should be accessed by only an authorized person. This concern is placated with the help of this technology—Blockchain (Fig. 1).

And once we add the concept of Artificial Intelligence into the concept of securing medical data in the blockchain, it will eventually become smarter and more secure by automatically realizing that this data is of concern and to be secured and which one needs to be discarded [11].

### 3.3 Technologies Used and Software

#### 3.3.1 Smart Contracts

Smart contracts are the brain of blockchain so is the most important component to be deployed along with blockchain to IoT devices. Specifically, smart contracts can be considered as scripts and are written in the form of conditional statements and if true actions will be triggered else not [12].

#### 3.3.2 Ethereum

Ethereum is a distributed computing platform, which is based on blockchain and is open source and public. It also features smart contract functionality [13]. It is actually a modified version of Nakamoto consensus through state transition, which



Fig. 1 Digital signature formation

is based on transaction. For example, Cryptocurrency like ether is generated by this blockchain platform. It is written in Go, C++, and Rust [1].

### 3.3.3 Python

Python is a high-level general-purpose programming language. It is utilized in both machine learning and blockchain applications because of its scalability, portability, robustness, powerful design, etc. Python is also very easy to implement as compared to the other programming languages. It is well equipped with the huge number of inbuilt libraries which can be directly implemented in the AI and Blockchain.

### 3.3.4 Decentralized Apps (DApps)

Ethereum is a distributed computing platform, which is based on blockchain and is open source and public. It also features smart contract functionality [13]. It is actually modified version of Nakamoto consensus through state transition which is based on transaction. Example Cryptocurrency like ether is generated by this blockchain platform. It is written in Go, C++ and Rust [1].

## 4 Working

First of all, the IoT devices containing sensors will be provided to each patient and the sensors installed in them will continuously monitor the health of the patient carrying that device. The reports of the monitored data will be sent to the doctor on his mobile device and also to the server through a GSM module. Moreover, the data could also be viewed on a website (as shown in Fig. 2). And also, any abrupt change in the normal behavior in the patient's health will also be reported with a warning notification to the doctor (Fig. 3).

But with so much of data stored on the cloud, there is a need to secure that humongous amount of data and prevent it from tampering. This virtue is done with help of blockchain. For this, first of all we will have to write a code for the blocks in programming language like Python. Here we will create a class Block with a Block number, data, a pointer to the next block, and a hash function of the block. Most importantly a block has the hash function of the previous block which makes a blockchain immutable. There is also a timestamp related to a block; this timestamp helps in synchronization of the blockchain [14]. Now once the data of a patient is stored in the form of a string on the block, it gets added to the node along with hash describing the location of the block. Now the smart contract will come into play by connecting it with blockchain which will maintain the privacy and security of the blockchain. The patient–client relation is secured by deploying smart contracts on blockchain. And the data generated by sensors such as Blood Pressure Level,



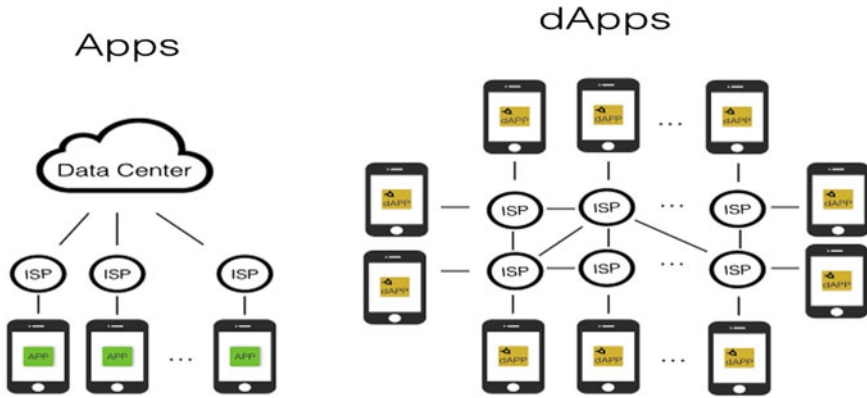


Fig. 2 Comparison between Apps and DApps

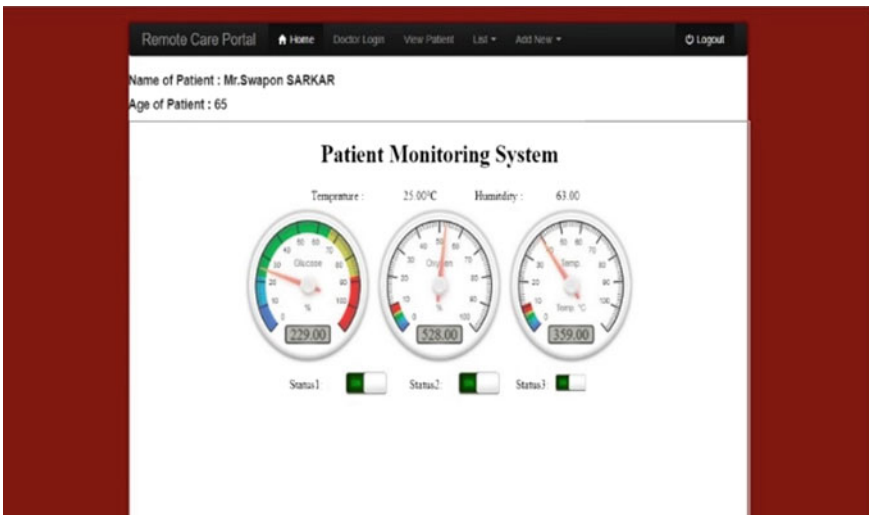


Fig. 3 Patient’s health data

ECG, etc., will be sent and stored in off-chain database like IPFS through gateways like mobiles and laptops. A hash included in blockchain which will be sent via notification will tell the location and will be sent through client, for example, clients of Ethereum–geth or PyEth.

As shown in Fig. 4, a patient’s health will be detected by sensors and a node is created by Arduino Uno and another node is created at a database IPFS. The patient at a remote location with sensors gets its checkups done and a block will be created with data and hash on it. The patient’s block will hold a private key. An authorized doctor will only be able to access that crypt block using its public key and no one else. Later on, using smart contracts the data will be stored on an off-chain database

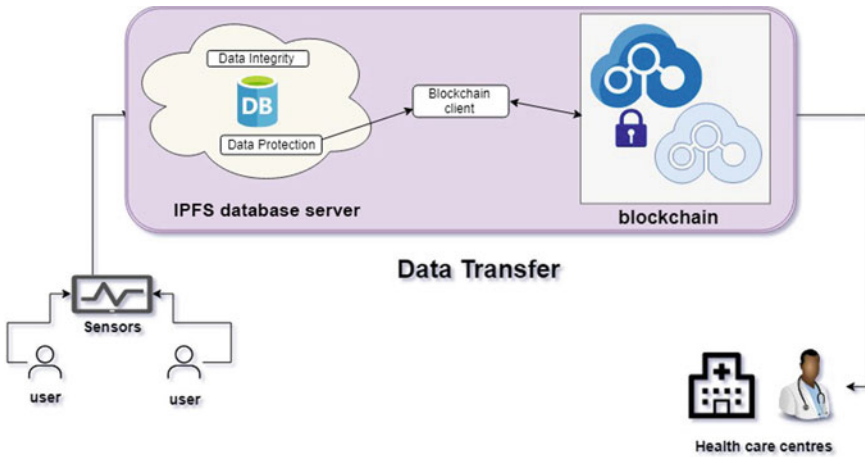


Fig. 4 Block diagram of the proposed model

IPFS and communication between doctor and patient will occur when the doctor will receive the notification about any abrupt change in the normal behavior or the reports of the patient.

Here, we are working on Ethereum environment we are working on the public domain. The authorized doctor will use his public key to access that information. After a new health report, a new block is added to the previous ledger using the hash of the previous block and a new hash is assigned to the new block. Since a cryptic language is used and a chain of blocks is made which is interconnected, now if a hacker will perform some tampering in any one of the blocks, he will have to change the data of in every block because a hash function is linked with each block. Moreover, the patient and the doctor will get to know about tampering. So now it becomes very tedious and almost impossible to tamper the data.

### 4.1 Observed Database Table

The Database Table will contain the patient id and name of the doctor assigned to the patient, respectively, along with the patient’s details such as date of report, temperature, ECG value, and Blood Pressure. Blockchain-related details like timestamp and hash history will also be found in the table (Fig. 5).

Patient ID	Doctor Assigned	Date Repoted	Timestamp	Temperature	Blood pressure	ECG	Hashed History
101	Dr. M	12/2/17	Tmp#1	102 F	85 bpm	value 1	hexa#1
409	Dr. Z	16/3/18	Tmp#2	97 F	110 bpm	value 2	hexa#2

Fig. 5 Observed database table

### 4.2 Flowchart of Working

As shown in figure, the real-time sensors will collect data such as ECG reading, pulse rating, blood pressure, and temperature data from their respective IoT sensors. These data will then be integrated into single unit based upon the time of the data retrieval and time stamp is allotted to the data unit in order to preserve date and time of the originated data. Since timestamps of the data are stored simultaneously in database, it prevents any possible malpractice with the data in future. With the help of hash function the data is then encrypted and stored in IFPS data server. The stored data is first extracted from IFPS data server and then decrypted in order to display it either on the patient’s account or the doctor’s account on server. The registered doctor can monitor patient’s health statistics anytime he feels like. If in case any aberration is found in patient’s health data, SMS will automatically be sent to the registered mobile number of doctor and relatives of patient (Fig. 6).

## 5 Conclusions and Future Scope

The research paper findings state that huge and critical data can be secured and be saved from any kind of tamper from the accomplice. So, we can conclude by denoting that Blockchain is the upcoming and the safest technology for security of data. All the necessary data can be saved on cloud and data analytics can be performed to observe some patterns on health problems based upon region, climate, time, number of patients with similar symptoms, etc. Patients who are willing to share their health records and medical data can be given some incentives in cash, which will inspire more patients to cooperate and aid in analysis. Further, the fees of prescribed doctors can also be paid through in-system online cash service which will also be protected by the blockchain. Also, the authorized family members and guardians can be provided facility through which they can access the website and monitor the patient’s condition from the distant location.

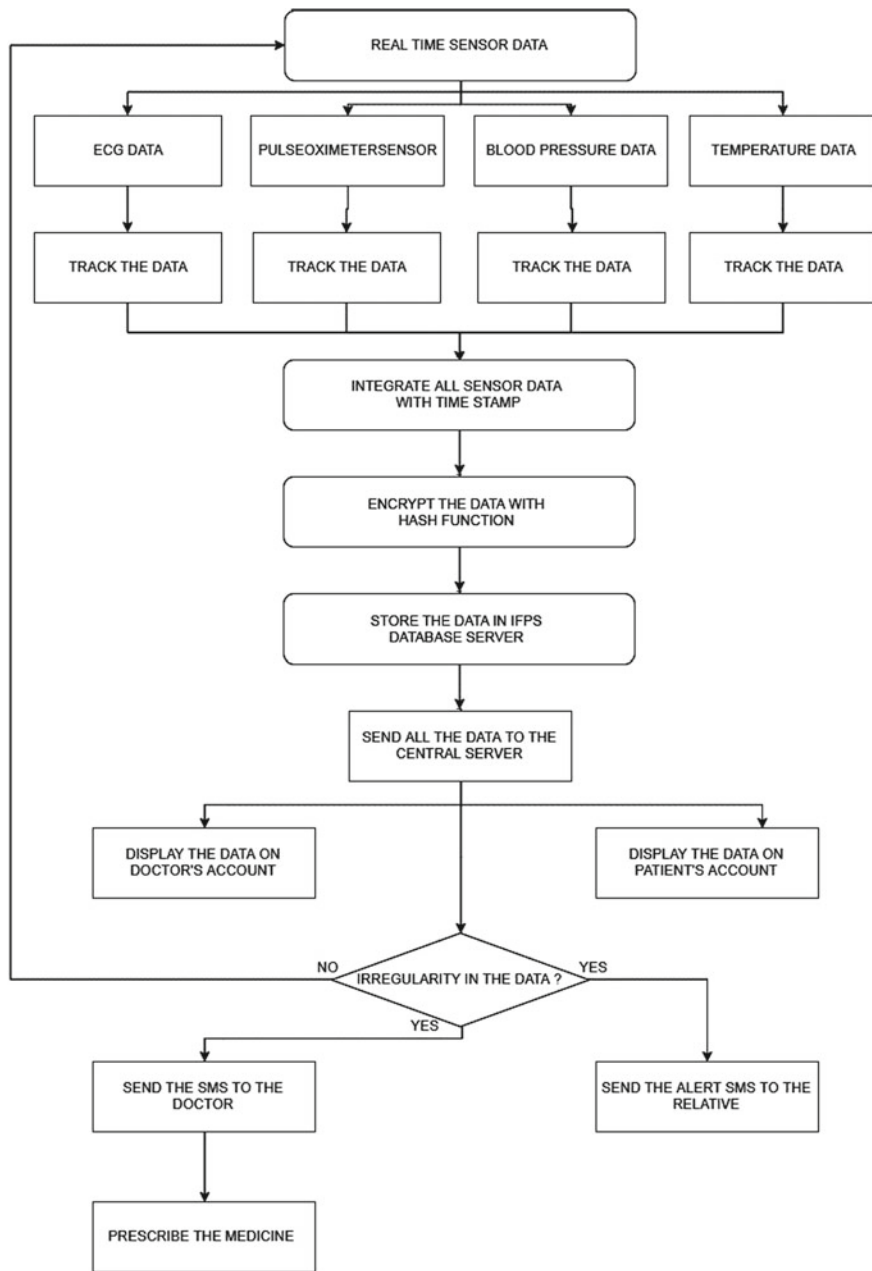


Fig. 6 Flowchart of the working of proposed mode

## References

1. Mohanta, B., Soumyashree, S. P., & Jena, D. (2018). *An overview of smart contract and use cases in blockchain technology*. <https://doi.org/10.1109/iccncnt.2018.8494045>.
2. Sun, Y., Zhang, R., Wang, X., Gao, K., & Liu, L. (2018). *A decentralizing attribute-based signature for healthcare blockchain*, pp. 1–9. <https://doi.org/10.1109/iccncn.2018.8487349>.
3. Le Nguyen, T. (2018). *Blockchain in healthcare: A new technology benefit for both patients and doctors*, pp. 1–6. <https://doi.org/10.23919/picmet.2018.8481969>.
4. Culver, K. (2016). *A whitepaper discussing how claims process can be improved*. [https://www.healthit.gov/sites/default/files/3-47-whitepaperblockchainforclaims\\_v10.pdf](https://www.healthit.gov/sites/default/files/3-47-whitepaperblockchainforclaims_v10.pdf).
5. Atlam, H., Alenezi, A., Alassafi, M., & Wills, G. (2018). Blockchain with Internet of things: Benefits, challenges and future directions. *International Journal of Intelligent Systems and Applications*, 10. <https://doi.org/10.5815/ijisa.2018.06.05>.
6. <https://www.youtube.com/user/CreatiiveCode>.
7. Hashemi, S. H., Faghri, F., Rauschy, P., & Campbell, R. H. (2016). World of empowered IoT users. In *2016 IEEE First International Conference on Internet-of-Things Design and Implementation (IoTDI)*. Das, M. L. (2015). Privacy and security challenges in Internet of things. In *Distributed Computing and Internet Technology* (pp. 33–48).
8. On Public and Private Blockchains. (2017). <https://blog.ethereum.org/2015/08/07/on-public-and-privateblockchains/>.
9. Rifi, N., Rachkidi, E., Agoulmine, N., & Taher, N. C. (2017). Towards using blockchain technology for eHealth data access management. In *2017 Fourth International Conference on Advances in Biomedical Engineering (ICABME)*. <https://doi.org/10.1109/icabme.2017.8167555>.
10. Blockchain: Opportunities for Health Care. (2016, August). Available: [deloitte.com/content/dam/Deloitte/us/Documents/publicsector/us-blockchain-opportunities-for-health-care.pdf](http://deloitte.com/content/dam/Deloitte/us/Documents/publicsector/us-blockchain-opportunities-for-health-care.pdf).
11. Ethereum Foundation. (2017). *Ethereum project*. Retrieved January 3, 2017, from <http://www.ethereum.org>.
12. Watanabe, H., Fujimura, S., Nakadaira, A., Miyazaki, Y., Akutsu, A., & Kishigami, J. (2016). Blockchain contract: Securing a blockchain applied to smart contracts. In *2016 IEEE International Conference on Consumer Electronics (ICCE)*. <https://doi.org/10.1109/icce.2016.7430693>.
13. Coblenz, M. (2017). Obsidian: A safer blockchain programming language. In *2017 IEEE/ACM 39th International Conference on Software Engineering Companion (ICSE-C)*. <https://doi.org/10.1109/icse-c.2017.150>.
14. Singh, M., Singh, A., & Kim, S. (2018). Blockchain: A game changer for securing IoT data. In *2018 IEEE 4th World Forum on Internet of Things (WF-IoT)*. <https://doi.org/10.1109/wf-iot.2018.8355182>.

# A Review of Big Data Challenges and Preserving Privacy in Big Data



Anil Sharma, Gurwinder Singh and Shabnum Rehman

**Abstract** We are living in an era where structured and unstructured data is produced, consumed and stored in enormous amount on frequent basis. Database transactions, social media, images, audios, videos etc. are the major sources responsible for generating big data in huge capacity and diversity. This usually consists of large volumes of complex and growing data sets with numerous self-regulating sources that are difficult to process with the conventional techniques of data management. Using big data mining, organizations are able to extract useful evidences from these large data sets. In spite of big data gains, there are numerous challenges also and among these challenges maintaining data privacy is the most important concern in big data mining applications since processing large scale of sensitive data sets such as health record, banking transaction records needs to be maintained in such a way that the private data should not be revealed to any unauthorized person. This paper provides a review of big data, challenges in big data mining and the privacy concern in big data.

**Keywords** Big data · Big data challenges · Big data technologies · Data mining · Data privacy

## 1 Introduction

The massive information in the organizations until now was just ordinary information maintained in the databases. All of a sudden, this gigantic information termed as big data got popular [1]. This refers to incredibly huge and complex data that become complicated to process using conventional applications [2]. Data can be both structured and unstructured, and may come from different sources such as transactions,

---

A. Sharma · G. Singh (✉) · S. Rehman  
School of Computer Applications, Lovely Professional University, Phagwara, Punjab, India  
e-mail: [gurwinder.11@gmail.com](mailto:gurwinder.11@gmail.com)

A. Sharma  
e-mail: [anil.19656@lpu.co.in](mailto:anil.19656@lpu.co.in)

S. Rehman  
e-mail: [shabnum148@gmail.com](mailto:shabnum148@gmail.com)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_7](https://doi.org/10.1007/978-981-15-0694-9_7)

social media, images, audio, and videos. It is not possible for traditional systems like SQL and RDBMS to deal with big data due to its scalability and complexity. Four 'V's, namely, Volume, Velocity, Variety, and Veracity represent the key features of Big Data [3].

The quality of data collected differs greatly and maintaining privacy and security of uncertain data is often a challenging task [4]. Data collected by the huge companies like Google, Facebook, and Twitter shows the need of big data management. Such organizations contain vast amount of data about individuals that need to be analyzed in order to find some valuable information [5]. Nowadays, the term big data is used universally in every field like finance, banking, and marketing where every day and every second data flows through workstations which are well managed and stored in order to find out the consumer behavior or the market trends to gain profitability [6].

The pace at which the variety and volume of data is produced makes very hard for the conventional data processing practices to cope up with these huge data sets [7]. These huge data sets require some advanced tools and technologies to deal efficiently during processing of large data [4]. Some of the solutions for the above-discussed issues are Hadoop, MapReduce, No SQL, Apache Spark, Hbase, Pig, Hive, Sqoop, and Oozie. Among these tools, Hadoop has the utmost role for handling big data [4, 7].

There can be different perspectives to look into the data to draw attention-grabbing patterns and further to gain the concealed awareness that can be helpful to forecast future drifts in the business market to gain sure returns. As the data mining plays the key role in the process of knowledge discovery, therefore it is usually denoted as Knowledge Discovery Databases (KDD) [8, 9]. If data is not analyzed or mined appropriately, it is worthless; therefore, for the markable growth of organizations it is better to analyze the data [10]. Data mining is used in every field like business, science, and engineering as data analysis has provided numerous benefits to the humanity as well as organizations [2]. Preserving privacy is one of the crucial elements of data mining, which means to ensure the security of individual information by avoiding any unauthorized access on it as it can result in adverse consequences. Thus, privacy is an important aspect that organizations need to maintain while mining results in order to not disclosing individual identity [11].

## 2 Privacy Concern in Big Data

Designing privacy enhancing techniques are getting a lot of attention over the period. Methodical, abstract, and legitimate significance of data privacy are always important highlights. There is always a serious issue with big data mining because it requires individual data in order to bring noteworthy effects [12]. Organizations collect large amount of data that contains individuals' specific information like in health records, financial transactional records, user name, contact, email, date of birth, credit card number refers to personal data [13]. Sometimes this data needs to be shared among

the third party at that time we need to maintain the privacy [14]. To secure data against unauthorized access is an important task for organizations [15].

Consider social sites like Facebook, Twitter, Google, Orkut a few individuals enthusiastically transfer their own personal details which may include audio, video, and images. This leads to the increase in cybercrimes, as it can be considered as serious problem especially among Facebook generation [16, 17].

When the data is processed, analyzed, and modified, it is the responsibility of security manager to maintain security at each level of architecture in order to preserve data at each level from malicious attacks, unwanted inference, or unauthorized access. Security may contain confidentiality, authenticity, and availability [18–20]. For any business, it is better to maintain privacy within an enterprise because identity of user if revealed will cause serious problems for the organization [21]. With the advancement and the internet usage, there is an increase in the rate of threats against the privacy [22]. Therefore, there are numerous strategies including encryption procedures and data structures, which maintain unawareness to patterns of accessing data and various privacy techniques used to change the information to make it more complex to link particular information records to particular people. To preserve privacy of big data, data mining techniques plays an important role.

### 3 Literature Review

The exponential growth of data results in varied complications for the organizations which bring some common challenges like Heterogeneity, Scalability, Timeliness, Infrastructure Fault, and Skill Requirements from Data Investigation and Storage viewpoints [10].

In [11], an algorithm for multilevel security using masking is recommended to pinpoint the complex stakes of big data. It allows Data collector to acquire data from the data source in scrambled and cleared form. Cleared data is directed to a database where stringent guidelines are followed, and only scrambled data is sent to data miner by data collector. Before extracting knowledge data miner connects present data to the sensitive data by applying decryption and if data matches, then it is provided to decision maker if not then it is returned back to the data miner for further improvement. The problem with this algorithm is every time data is matched to sensitive data stored in the database which is very time consuming.

In [15], cryptographic algorithm is proposed to protect the data by converting plain text into ciphertext using encryption schemes. The model used in algorithm consists of three layers: Secret, Authorized, and Public layer. At Secret layer, data is encrypted and a digital mark is attached to the data and is accessible to only authorized persons. At the intermediate layer, authorized persons having the private key can decrypt the data and perform data mining techniques. In public layer, the conclusions drawn afterward the data mining procedure are observed and also allowing authorized person to view self-information. The only problem is that less sensitive data that can be fruitful in the analysis of big data is also encrypted and is not accessible.



In [21], some of the factual issues interrelated to big data processing, storage, and management are emphasized along with various challenges that might be encountered in future due to the epidemic growth of data. Besides velocity, volume, and variety of data, complexity is an additional feature of big data and moreover handling of big data is the real issue highlighted by the author.

In [23], some insights about big data issues, challenges, and tools alongside basic concepts and properties of big data like velocity, volume, heterogeneity, are talked about. Furthermore, different sources from which data is generated are examined. Big data has a huge significance in different activities like community media, sensor information, and log storage and risk analysis.

In [24], problems related to privacy are cited with the recommendation to make use of organization authentication for big data using MapReduce, processing of data and privacy preserving. Integration of MapReduce, if used for analyzing data may provide better privacy.

In [25], a data-driven big data processing model is proposed, from viewpoint of data mining, which encompasses demand obsessed collection of information bases, privacy, user concern modeling, mining and exploration, and security aspects. It also introduces a three-tier architecture framework, where first tier is focusing on accessing data and arithmetic computing, and second tier is concentrating on the user privacy issues and the third tier is aiming towards challenges faced while mining the complex and dynamic data.

In [26], a technique, K-anonymity, is introduced in which every record is alike to at least other  $k-1$  other records on the possibly recognized variables. K-anonymity can be achieved using generalization and overpowering (replaces original value by some special character like \*). The only issue with this technique is that it doesn't give attention to the links between the sensitive attributes so there is still outflow of sensitive data.

In [27], the authors proposed a generalization algorithm generally called as bottom-up approach in order to deal with the scalability of data. The structure for generalization can be obtained by making a tree of user's original data set and various operations can be performed on specific ranges. This way of anonymizing data may be considered as efficient because generalization compresses the user's data as data increases. Identifying the best generalization is the key to climb up the hierarchy at each iteration.

In [28], a homomorphic technique is developed which is basically a form of encryption that allows performing some specific computations on ciphertext and encrypted results are obtained. The decrypted results are then matched to the results of operations that are performed on plain text. This approach is useful to deal with the entrusted party because neither the input is unveiled nor the internal state of the encrypted data.

In [29], top-down specialization approach is introduced that provides security, and preserves sensitive data of the user by partitioning the large data sets into two phases; in first phase data is anonymized and intermediate results are created, while in the second phase the first phase results are combined to get the ultimate outcome. The only problem with this approach is that if the data set is too large it becomes

difficult to apply anonymization to the data and there remains fair of privacy losses while portioning the data.

In [30], a method for securing two-party high multidimensional private data is introduced called data mashup technique. This technique generally mashes up the data on users end before it is sent to the third party. Only the ordinary data is exposed to third party, and the sensitive data is hidden by performing encryption before it is revealed to the other party. The issue with this technique is mashing up large data sets will require a lot of time.

In [31], a technique called differential privacy is mentioned as method that doesn't allow clients to have access to the database. It is totally opposed to anonymization, as there is no need to modify the data but an interface exists that calculates results and adds distortion to the results, and after this the results are displayed. The only aim of this technique is to shrink the possibilities of individual recognition while querying the data. One problem with this method is that an analyst should know the query before using it.

In [32], a proxy re-encryption technique involves only sharing of ciphertext securely over multiple times. Neither the message and sender's identity nor the receiver's identity is disclosed. Basically, it follows an encryption scheme that allows converting the ciphertext of particular key into an encryption of the same message by using another separate key.

In [33], some of the detailed technologies have been discussed like generalization, bucketization, and multiset-based generalization, one attribute per column, slicing, and slicing with suppression. By using these techniques a different level of privacy can be achieved. Generalization technique is difficult to apply on high-dimensional data. Bucketization fails to maintain the membership disclosure, so they have mentioned slicing technique that can be used to overcome the above problems.

In [34], slicing technique, which is basically an anonymizing technique can partition data vertically as well as horizontally. In vertical partitioning, attributes that highly correlate to each other are clustered into column. In horizontal partitioning, column values are sorted randomly so that no column values can be linked. Slicing is mainly used to not only to interrupt the relationship across columns, but to ensure the bond between each column. To deal with the high-dimensional data, slicing technique is the best approach.

In [35], hybrid technique is proposed by combining randomization and generalization. First, data randomization is performed and after that generalization method is applied to the randomized data. The technique provides better accuracy by reconstructing original data without any loss of information. In [36], an output perturbation privacy maintaining approach is proposed with help of differential privacy to improve accuracy of query processing and reducing the possibility of leakage of privacy.

## 4 Findings

Challenges faced by various privacy techniques are delineated in Table 1.

Table 2 shows a comparative analysis of some of the privacy-preserving techniques based on parameters linkage property, information loss, type of data, and privacy preserved.

The analysis results in that no single method is reliable in all spheres. Each method performs in a different way depending on the size of data and the type of application.

**Table 1** Privacy techniques and challenges

Techniques	Challenges
Slicing [2]	Mostly, the attributes are grouped randomly which is not efficient It's not clear how attribute disclosure is preserved Utility of data is lost because of fake tuples
Cryptographic technique [15]	Difficult to apply for large databases Difficult to scale when more events are involved Non-sensitive data is also encrypted that can be useful for analytics
Differential privacy [22]	High computation complexity No preservation of data truthfulness at the record level
K-anonymity [26]	Gives no consideration of the links between sensitive data Not able to protect against attacks based on background knowledge Not applicable for high-dimensional data
Anonymization through generalization [27]	Causes loss of information Not ready to protect attribute correlations Each attribute is generalized separately To climb up the hierarchy, each iteration needs to recognize the best generalization
Homomorphic encryption [28]	Computational overhead increased Not applicable for large datasets
Top-down specialization approach [29]	Loss of privacy Leads to its inadequacy in handling large-scale data sets
Data mashup technique [31]	Mashing large scale of data requires a lot of time Mashing of data may cause a loss of accuracy
Bucketization [33]	Can't intercept attribute: membership disclosure Essential to Issues Quasi Identifiers values in their original form Needs clear split-up between quasi-identifiers and sensitive attributes

**Table 2** Comparison of different privacy techniques

Techniques	Parameters			
	Linkage property	Information loss	Type of data	Privacy preserved
Slicing technique [2]	Very low	Low	High dimensional	High
Cryptographic technique [15]	Low	Low	Micro data	High
Differential privacy [22]	Low	Low	Micro data	High
K-anonymization [26]	High	Low	Micro data	Low
Anonymization through generalization [27]	High	Very high	Micro data	Low
Homomorphic encryption [28]	Low	Low	Micro data	High
Top-down specialization technique [29]	Low	High	Micro data	Low
Data mashup technique [31]	Low	High	High dimensional	High
Proxy re-encryption [32]	Low	Low	Micro data	High
Bucketization [33]	High	Low	Micro data	Low
Hybrid approach [35]	Low	High	High dimensional	High

## 5 Conclusion and Future Work

Big data refers to the complex and huge data sets and big data mining is a process of discovering unknown patterns from big data. With the rising and quickly growing data, things are varying in the business environment. Big data is fetching the hottest ultimate edge for data research and for many business applications. Companies are currently using big data analysis to forecast the upcoming trends so that enormous value can be produced out of it. Big data mining is an emerging research area; a constrained work has been done on it so far. Authors believe that a much of work needs to be done in order to overcome its challenges like heterogeneity, scalability, infrastructure faults, timeliness, and privacy. More precisely, authors pointed the privacy challenges of big data mining. The extreme volume, velocity, and variety of data is creating problem for most of the organizations because they are not able to protect such volumes of data against different attacks. In big data mining it is not possible to carry out the operations without compromising the privacy. Business organizations hold sensitive information about their clients and this information is considered as a big asset to them. To safeguard this information against unauthorized access, few techniques are proposed in the literature but have limitations. So, the authors believe that more such techniques and mechanism need to be developed that will help in preserving privacy during data analysis process, for the reason that if privacy about an individual is violated it may have catastrophic significance on someone’s life.

## References

1. Ahsan, U., & Bais, A. A. (2016). Review on big data analysis and Internet of things. In *2016 IEEE 13th International Conference on Mobile Ad Hoc and Sensor Systems* (pp. 325–330). IEEE. <https://doi.org/10.1109/mass.2016.38>.
2. Chakraborty, N., & Gonnade, S. (2014). Big data and big data mining: Study of approaches, issues and future scope. *International Journal of Engineering Trends and Technology*, *18*, 221–223.
3. Sawant, P. G., & Desai, B. L. (2015). Big data mining: Challenges and opportunities to forecast future scenario. *International Journal of Innovative Technology and Exploring Engineering*, *3*, 5228–5232.
4. Hashem, I. A. T., et al. (2015). The rise of ‘big data’ on cloud computing: Review and open research issues. *Information Systems*, *47*, 98–115.
5. Kalbandi, I., & Anuradha, J. A. (2015). Brief introduction on big data 5Vs characteristics and Hadoop technology. In *Procedia Computer Science 48: International Conference on Computer, Communication and Convergence (ICCC 2015)* (Vol. 48, pp. 319–324).
6. Tole, A. A. (2013). Big data challenges. *Database Systems Journal*, *IV*, 31–40.
7. Sagiroglu, S., & Sinanc, D. (2013). Big data : A review. In *2013 International Conference on Collaboration Technologies and Systems (CTS)* (pp. 42–47). IEEE.
8. Dev, H., Sen, T., Basak, M., & Ali, M. E. (2013). An approach to protect the privacy of cloud data from data mining based attacks. In *2012 SC Companion: High Performance Computing, Networking Storage and Analysis* (pp. 1106–1115). IEEE. <https://doi.org/10.1109/sc.companion.2012.133>.
9. Singh, D. K., & Swaroop, V. (2013). Data security and privacy in data mining: Research issues & preparation. *International Journal of Computer Trends & Technology*, *4*, 194–200.
10. Sameer, A. (2016). Big data and data mining a study of (characteristics, factory work, security threats and solution for big data, data mining architecture, challenges & solutions with big data). In *Advancing Web Paging Techniques* (pp. 1–XXVI). <https://doi.org/10.13140/rg.2.1.3238.9525>.
11. Choopa, M. R. (2015). Data mining and security in big data. *International Journal of Advanced Research in Computer Engineering and Technology*, *4*, 1065–1069.
12. Hbib, L., & Barka, H. (2016). Big data: Framework and issues. In *2016 International Conference on Electrical and Information Technologies (ICEIT)* (p. 6). IEEE.
13. Nargundi, S. M., & Phalnikar, R. (2013). Data DE-identification tool for privacy preserving data mining. *International Journal of Computer Science Engineering and Information Technology Research*, *3*, 267–276.
14. Sriyayanthi, S., & Sethukkarasi, R. (2017). A comprehensive survey on privacy preserving big data mining. *International Journal of Computer Applications Technology and Research*, *6*, 79–86.
15. Hussain, N. I., Choudhury, B., & Rakshit, S. (2014). A novel method for preserving privacy in big-data mining. *International Journal of Computers and Applications*, *103*, 21–25.
16. Kaushik, M., & Jain, A. (2014). Challenges to big data security and privacy. *International Journal of Computing Science and Information Technology*, *5*, 3042–3043.
17. Smith, M., Szongott, C., Henne, B., & Voigt, G. Von. (2013). Big data privacy issues in public social media. In *2012 6th IEEE International Conference on Digital Ecosystems and Technologies (DEST)* (p. 6). IEEE.
18. Geethakumari, G., & Srivatsava, A. (2012). Big data analysis for implementation of enterprise data security. *International Journal of Computer Science, Information Technology, & Security*, *2*, 742–746.
19. Jaseena, K. U., & David, J. M. (2014). Issues, challenges, and solutions: Big data mining. In *Sixth International Conference on Networks & Communications* (pp. 131–140).
20. Kim, S., & Lee, I. (2015). Data block management scheme based on secret sharing for HDFS. In *10th International Conference on Broadband and Wireless Computing, Communication and Applications Data* (pp. 51–56). IEEE. <https://doi.org/10.1109/bwcca.2015.70>.

21. Kaisler, S., Armour, F., Espinosa, J. A., & Money, W. (2013). Big data : Issues and challenges moving forward. In *2013 46th Hawaii International Conference on System Sciences* (pp. 995–1004). IEEE. <https://doi.org/10.1109/hicss.2013.645>.
22. Xu, L. E. I., Jiang, C., Wang, J., Yuan, J., & Ren, Y. (2014). Information security in big data : Privacy and data mining. *IEEE Access*, 2, 1149–1176.
23. Katal, A., Wazid, M., & Goudar, R. H. (2013). Big data : Issues, challenges, tools and good practices. In *2013 Sixth International Conference on Contemporary Computing (IC3)* (pp. 404–409). IEEE.
24. Vennila, S., & Priyadarshini, J. (2015). Scalable privacy preservation in big data a survey. In *2nd International Symposium on Big Data and Cloud Computing (ISBCC'15)*. Vijayakumar, V., & Neelanarayanan, V. (Eds.). (2015). *Procedia Computer Science*, 50, 369–373 (Elsevier B.V).
25. Wu, X., Zhu, X., & Member, S. (2014). Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering*, 26, 97–107.
26. Salini, S., Kumar, S. V., & Neevan, R. (2015). Survey on data privacy in big data with K-anonymity. *International Journal of Innovation and Research Computer Communication*, 3, 3765–3771.
27. Balusamy, M. (2014). Data anonymization through generalization using map reduce on cloud. In *2014 IEEE International Conference on Computer Communication and Systems (ICCCS '14)* (pp. 39–42). IEEE.
28. Sangeetha, M., Anishprabu, P., & Shanmathi, S. Homomorphic Encryption Schema for Privacy Preserving Mining of Association Rules. *International Journal of Innovation Research Science Engineering*.
29. Fung, B. C. M., Wang, K., & Yu, P. S. (2005). Top-down specialization for information and privacy preservation. In *ICDE '05 Proceedings of the 21st International Conference on Data Engineering* (pp. 205–2016). IEEE.
30. Sridhar, I., & Jacob, P. (2014). Secure two party high dimensional private data using data mash up. *International Journal of Computer Science and Information Technologies*, 5, 644–645.
31. Gosain, A., & Chugh, N. (2014). Privacy preservation in big data. *International Journal of Computers and Applications*, 100, 44–47.
32. Liang, K., Susilo, W., & Liu, J. K. (2015). Privacy-preserving ciphertext multi-sharing control for big data storage. *IEEE Transactions on Information Forensics and Security*, 10, 1–11.
33. Kaur, P. C., Ghorpade, T., & Mane, V. (2016). Analysis of data security by using anonymization techniques. In *2016 6th International Conference—Cloud System and Big Data Engineering (Confluence)* (pp. 287–293). IEEE.
34. Rodiya, K., & Gill, P. (2015). A review on anonymization techniques for privacy preserving data publishing. *International Journal of Engineering Research and Technology*, 4, 228–231.
35. Lohiya, S., & Ragha, L. (2012). Privacy preserving in data mining using hybrid approach. In *2012 Fourth International Conference on Computational Intelligence and Communication Networks* (pp. 743–746). IEEE. <https://doi.org/10.1109/cicn.2012.166>.
36. Du, M., Wang, K., Chen, Y., Wang, X., & Sun, Y. (2018). Big data privacy preserving in multi-access edge computing for heterogeneous Internet of things. *IEEE Communications Magazine*, 56, 62–67.

# Dual-Layer DNA-Encoding–Decoding Operation Based Image Encryption Using One-Dimensional Chaotic Map



K. Abhimanyu Kumar Patro , M. Prasanth Jagapathi Babu,  
K. Pavan Kumar and Bibhudendra Acharya 

**Abstract** This paper describes a technique that encrypts images in the form of DNA sequences using PWLCM system, i.e., Piecewise Linear Chaotic Map. This method has two times DNA-encoding–decoding operations along with DNA-permutation and DNA-diffusion operations to get the cipher image. On comparison with other processes, the advantage of this algorithm is easy to compute but confuses a cryptanalyst a lot. Apart from that dual-layer DNA-encoding–decoding processes in the algorithm result in good encryption outputs. When outputs are subjected to different security analysis to find out the strength of the algorithm, results with encrypted images with higher values of UACI, NPCR, key space, information entropy, and good correlation coefficient. This results in strong resistivity toward widely used attacks.

**Keywords** Image encryption · One-dimensional chaotic map · Security · DNA operations · SHA-256

## 1 Introduction

In a communication system, people exchange data in the form of text, images, audios, and videos. Most commonly used are images and securing them is important. To achieve this, different conventional encryption processes such as Rivest–Shamir–Adleman (RSA), DES, Triple-DES (3-DES), and Advanced Encryption Standard (AES) [1, 2] are used but these encryption processes are inefficient to encrypt images

---

K. A. K. Patro · M. Prasanth Jagapathi Babu · K. Pavan Kumar · B. Acharya (✉)  
Department of Electronics and Telecommunication Engineering, National Institute of Technology  
Raipur, Raipur, India  
e-mail: [bacharya.etc@nitrr.ac.in](mailto:bacharya.etc@nitrr.ac.in)

K. A. K. Patro  
e-mail: [kpatro.phd2016.etc@nitrr.ac.in](mailto:kpatro.phd2016.etc@nitrr.ac.in)

M. Prasanth Jagapathi Babu  
e-mail: [jagapathi.matta@gmail.com](mailto:jagapathi.matta@gmail.com)

K. Pavan Kumar  
e-mail: [pavan.19.kpk@gmail.com](mailto:pavan.19.kpk@gmail.com)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_8](https://doi.org/10.1007/978-981-15-0694-9_8)

and secure them since they have bulky data, more redundancy, and more correlation of adjacent pixels [3–6].

To overcome these problems, researchers have suggested different methods to achieve a highly efficient way to encrypt these images and protect them from common attacks and one of such processes is called ‘Chaos-based encryption technique’ by Matthews, for the first time [7]. This technique helps in constructing secure cryptosystems. Since chaos maps are much sensitive to their initial conditions, hence they can be used to permute the image pixels in such a way that are not easy to decrypt with a small change of a key. Chaotic maps are classified into two types. They are one-dimensional (1D) and high-dimensional. The latter is also known as hyperchaotic map [8]. The 1D maps are highly efficient, simpler in structure, easy to implement in both software and hardware, and have less computational cost compared to high-dimensional chaotic maps [8, 9]. So, 1D maps are preferred while implementing image encryption algorithms.

These days a new encryption process which uses DNA computing is very popular in the field of cryptography. When compared to other encryption processes it has unique distinctions which make it more advantageous than other processes. They are less space to store, parallelism, minimal power requirement, etc.

But it has certain disadvantages such as the requirement of large number of steps for DNA processing, no universal property of solving problems and since the process is not an automated one it requires human attention, etc. [10]. So, encrypting images using DNA alone is not a good idea. To overcome these problems, we will encrypt images with DNA combined with chaotic maps. This ensures that the problems in security are removed [11]. Many image encryption techniques [12–14] have used DNA sequence operations along with 1D chaotic maps.

In this work, we have used DNA sequence operations along with 1D chaotic maps to perform image encryption. Many of the DNA based image encryption techniques [12–14] have used one-stage of DNA-encoding–decoding operation to change the order of the pixels of images, which somehow gives chance for a cryptanalyst to break down the algorithm. This paper, however, shuffles the pixels by encoding them followed by decoding with two different keys respectively, for two times hence it is called dual-layer encoding and decoding process which enhances the shuffling of pixels that results in higher security.

This paper contributes the following.

- Two times encoding and decoding is performed in this algorithm to shuffle the image pixels and make it hard to decrypt.
- To increase the confusion and to achieve a high diffusion rate among the pixels, a PWLCM map is used.
- Hash-based keys are used to resist Known-plaintext Attack (KPA) and Chosen-plaintext Attack (CPA) attacks.

The innovation process is described as mentioned. Section 2 explains the basics of the theme. The design flow for the encryption process is discussed in Sect. 3. Security analysis and simulation results are discussed in Sect. 4. In Sect. 5 conclusion is given.



## 2 Preliminaries

### 2.1 PWLCM

PWLCM stands for piecewise linear chaotic map. Because of being less impacted by external disturbances, it is used mostly in encryption processes [15–19]. It is formulated as

$$g_{n+1} = \begin{cases} \frac{g_n}{n} & \text{if } 0 \leq g_n < n \\ \frac{g_n - n}{0.5 - n} & \text{if } n \leq g_n < 0.5 \\ (1 - g_n) & \text{if } 0.5 \leq g_n < 1 \end{cases} \quad (1)$$

where,  $g_n \in [0, 1]$  is initial value and  $n \in (0, 0.5)$  is the control parameter of PWLCM system.

### 2.2 DNA Sequence Operations

Four different nitrogenous bases are present in the DNA: ‘A’, ‘T’, ‘C’, and ‘G’, these bases won’t pair off in a random manner. ‘T’ combines only with ‘A’, and ‘G’ only pairs along ‘C’. From this, we can infer that the ‘A’ and ‘T’ bases are always harmonious to each other and ‘G’ and ‘C’ bases are also harmonious to one another [20, 21].

Similarly, in a binary system, we know that 10 and 00 are complementary with 01 and 11. By using this similarity that exists between the binary system and DNA, DNA cryptography has become an efficient encryption technique that is capable of dealing with the binary system in encryption. Using the binary system and DNA, the DNA encoding and decoding rules are mentioned in Table 1 & Rule-1 based DNA-XOR is shown in Table 2.

**Table 1** Encoding and decoding DNA rules

Rule	A	T	G	C
1	01	10	00	11
2	10	01	00	11
3	01	10	11	00
4	10	01	11	00
5	11	00	10	01
6	11	00	01	10
7	00	11	10	01
8	00	11	01	10

**Table 2** Rule-1 based DNA-XOR operation

XOR	A	T	G	C
A	A	T	G	C
T	T	A	C	G
G	G	C	A	T
C	C	G	T	A

### 3 Proposed Methodology

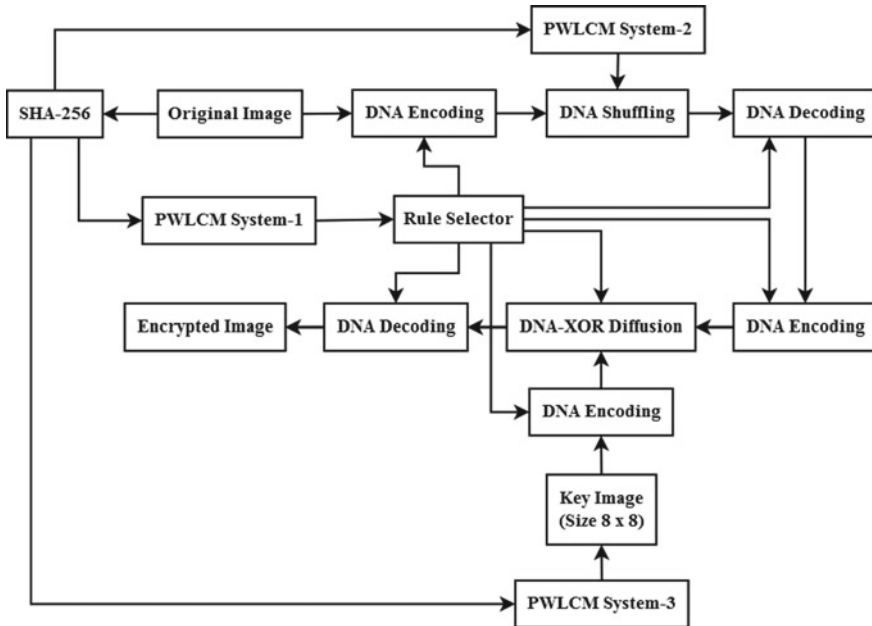
Figure 1 demonstrates the proposed methodology for image encryption. The steps are as below.

**Step 1:** Input an image ‘*I*’ of dimension “*M* × *N*”.

**Step 2:** Use the SHA-256 hash algorithm, which produces 64-hex values, denoted by

$$\text{hash} = h_1, h_2, h_3, \dots, h_{63}, h_{64} \tag{2}$$

**Step 3:** Obtain PWLCM System-1 based keys.



**Fig. 1** Proposed cryptosystem encryption block diagram

$$\begin{cases} pk11(1) = pk1(1) - \left( \left( \frac{\text{sum}(h_1 : h_{10})/10^{15}}{\text{ceil}(\text{sum}(h_1 : h_{10})/10^{15})} \right) - \right) / 10^2 \\ mue11 = mue1 - \left( \left( \frac{\text{sum}(h_{11} : h_{20})/10^{15}}{\text{ceil}(\text{sum}(h_{11} : h_{20})/10^{15})} \right) - \right) / 10^2 \end{cases} \quad (3)$$

where, ' $pk1(1)$ ' is the given initial value and ' $pk11(1)$ ' is the generated initial value of PWLCM System-1. Similarly, ' $mue1$ ' and ' $mue11$ ' are the given and produced system parameters of PWLCM System-1, respectively.

**Step 4:** Repeat PWLCM System-1 of Eq. (1) for 1008 rounds. Let the iterated PWLCM System-1 sequence is represented by ' $pk11$ '. In ' $pk11$ ', eliminate the first 1000 iterations to remove the transit effects. The newly obtained PWLCM System-1 sequence

$$pk11 = \{pk11(1), pk11(2), \dots, pk11(8)\} \quad (4)$$

**Step 5:** Classify the chaotic sequence 'system parameters of PWLCM System-1  $pk11$ ' in descending order. The ' $pk11$ ' sorting function is

$$[pk11\text{sort}, pk11\text{index}] = \text{sort}(pk11, \text{'descend'}) \quad (5)$$

where, the sorting sequence is denoted by ' $pk11\text{sort}$ ' and the indexing sequence is denoted by ' $pk11\text{index}$ '.

**Step 6:** Using the first index value of ' $pk11\text{index}$ ' (termed as DNA rule), encode the pixels of the image ' $I$ '. Let the encoded image be ' $IE$ '.

**Step 7:** Obtain PWLCM System-2 based keys.

$$\begin{cases} pk21(1) = pk2(1) - \left( \left( \frac{\text{sum}(h_{21} : h_{31})/10^{15}}{\text{ceil}(\text{sum}(h_{21} : h_{31})/10^{15})} \right) - \right) / 10^2 \\ mue21 = mue2 - \left( \left( \frac{\text{sum}(h_{32} : h_{42})/10^{15}}{\text{ceil}(\text{sum}(h_{32} : h_{42})/10^{15})} \right) - \right) / 10^2 \end{cases} \quad (6)$$

where, ' $pk2(1)$ ' is the given initial value and ' $pk21(1)$ ' is the generated initial value of PWLCM System-2. In the exactly same way, ' $mue2$ ' and ' $mue21$ ' are also the given and generated system parameters of PWLCM System-2, respectively.

**Step 8:** Repeat PWLCM System-2 of Eq. (1)  $M \times N \times 4$  times. The iterated PWLCM System-2 sequence is expressed as ' $pk21$ '. The newly formed PWLCM System-2 sequence is Similarly, ' $mue1$ ' and ' $mue11$ '

$$pk21 = \{pk21(1), pk21(2), \dots, pk21(M \times N \times 4)\} \quad (7)$$

**Step 9:** The chaotic sequence ' $pk21$ ' is sorted in descending order. The ' $pk21$ ' sorting function is

$$[pk21\text{sort}, pk21\text{index}] = \text{sort}(pk21, \text{'descend'}) \quad (8)$$

where, the sorting sequence is denoted as ' $pk21sort$ ' and the indexing sequence is denoted as ' $pk21index$ '.

**Step 10:** Using the indexing sequence ' $pk21index$ ', shuffle the DNA encoded pixels of ' $IE$ '. Let the shuffled output is denoted as ' $IEs$ '.

**Step 11:** Using the second index value of ' $pk11index$ ' (termed as DNA rule), decode the image ' $IEs$ '. Let the decoded image is represented as ' $IED$ '.

**Step 12:** Using the third index value of ' $pk11index$ ' (termed as DNA rule), encode the pixels of the image ' $IED$ '. Let the encoded image is represented as ' $IEDE$ '.

**Step 13:** Obtain PWLCM System-3 based keys.

$$\begin{cases} pk31(1) = pk3(1) - \left( \left( \frac{\text{sum}(h_{43} : h_{53})/10^{15}}{\text{ceil}(\text{sum}(h_{43} : h_{53})/10^{15})} \right) - \right) / 10^2 \\ \text{mue31} = \text{mue3} - \left( \left( \frac{\text{sum}(h_{54} : h_{64})/10^{15}}{\text{ceil}(\text{sum}(h_{54} : h_{64})/10^{15})} \right) - \right) / 10^2 \end{cases} \quad (9)$$

where, ' $pk3(1)$ ' is the given initial value and ' $pk31(1)$ ' is the generated an initial value of PWLCM System-3. However, ' $\text{mue3}$ ' is the given and ' $\text{mue31}$ ' is the generated system parameters of PWLCM System-3.

**Step 14:** Repeat PWLCM System-3 of Eq. (1) for  $8 \times 8$  rounds. Let the iterated PWLCM System-3 sequence is denoted as ' $pk31$ '. The newly formed PWLCM System-3 sequence is

$$pk31 = \{pk31(1), pk31(2), \dots, pk31(8 \times 8)\} \quad (10)$$

Using the sequence ' $pk31$ ', generate a key image ' $KeyI$ ' of size  $8 \times 8$ .

**Step 15:** Using the fourth index value of ' $pk11index$ ' (termed as DNA rule), encode the key image ' $KeyI$ '. Let the encoded image is denoted as ' $KeyIE$ ' of size  $1 \times 256$ .

**Step 16:** Split up the image ' $IEDE$ ' into small pieces of size  $1 \times 256$ .

**Step 17:** Using the fifth index value of ' $pk11index$ ' (termed as DNA rule), perform DNA-XOR diffusion for ' $KeyIE$ ' and the blocks of ' $IEDE$ '. The diffusion operation is as follows.

First diffusion is between ' $KeyIE$ ' and the first block of ' $IEDE$ '. The output is DNA-XOR'ed with the second block of ' $IEDE$ '. The DNA-diffusion process is continuing until the diffusion of all the blocks of ' $IEDE$ '.

Add every diffused block to form a diffused image ' $IEDEd$ '.

**Step 18:** Using the sixth index value of ' $pk11index$ ' (termed as DNA rule), decode the image ' $IEDEd$ '. Let the decoded image is denoted as ' $IEDED$ '. The decoded image is the cipher image of the proposed cryptosystem.

Doing the above steps in reverse order, we will get the decrypted image.

## 4 Security Analysis and Computer Simulations

Innovated scheme is checked by taking two images. The images taken are grayscale images. They are “Cameraman.tif” of size  $(256 \times 256)$  and “Lena.tif” of size  $(512 \times 512)$ . The computer simulations is done on PC having i3 processor (2.00 GHz), 64-bit windows OS, 4 GB RAM, and using MATLAB version R2016a. The different keys that are used, are in Table 3. Figure 2 illustrates the simulation results of the suggested scheme. By observing the outputs, one can infer that this method produces good encryption results. All the images required for the simulation process are taken from USC-SIPI image database [22].

### 4.1 Key Space Analysis

Different keys employed for the process are

- Keys of all the three PWLCM systems.
- 256-bits of SHA-256 hash algorithm.

According to IEEE,  $10^{-15}$  is data representation [23] standard for floating points. So, the suggested method uses a secret key with the accuracy of  $10^{-15}$ . To defend against the common attacks SHA algorithm produces key space of  $2^{128}$ .

Therefore, the key space of our innovation is  $(10^{15} \times 10^{15}) \times (10^{15} \times 10^{15}) \times (10^{15} \times 10^{15}) \times 2^{128} = 1.963 \times 2^{426}$  which is a larger value when compared to  $2^{128}$  [24]. The key space comparison results are as shown in Table 4. They show that the mentioned process possesses strong resistance toward brute-force attack than the processes in [25, 26].

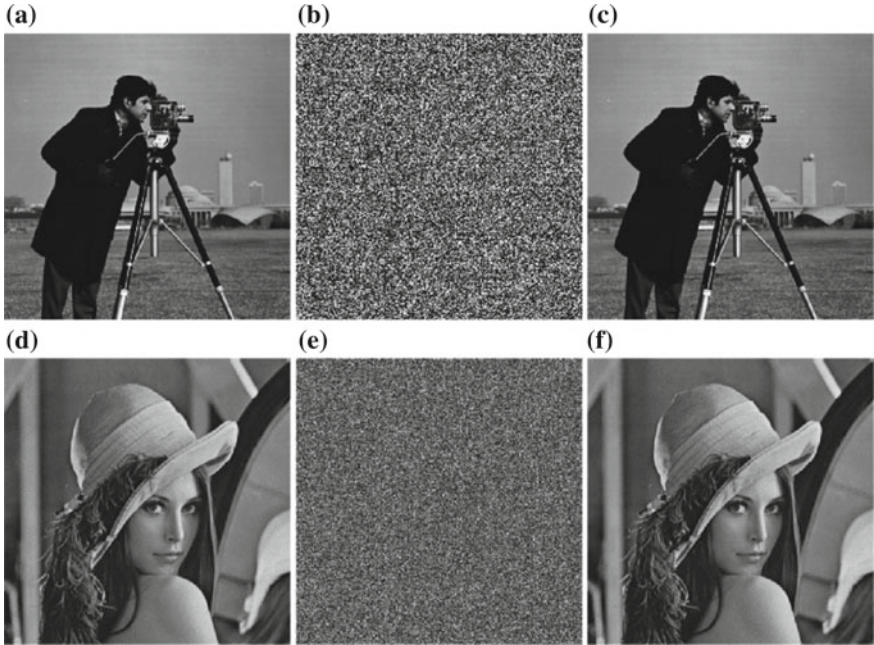
### 4.2 Statistical Attack Analysis

#### 4.2.1 Histogram Analysis

It is a graphical display of allotment of data. Encrypted images should have a homogenous histogram, so it is not possible for a cryptanalyst to retrieve the information

**Table 3** Original key values of the suggested scheme

Map used	Original keys	
	Initial values	System parameters
PWLCM System-1	$pk1(1) = 0.275648900231572$	$mue1 = 0.347823654894159$
PWLCM System-2	$pk2(1) = 0.275648900232379$	$mue2 = 0.347823654895732$
PWLCM System-3	$pk3(1) = 0.275648900231864$	$mue3 = 0.347823654892182$



**Fig. 2** Simulation results: **a** “Cameraman Plain Image” and **d** “Lena Plain Image”. **b** “Cameraman Cipher Image” and **e** “Lena Cipher Image”. **c** “Cameraman Decrypted Image” and **f** “Lena Decrypted Image”

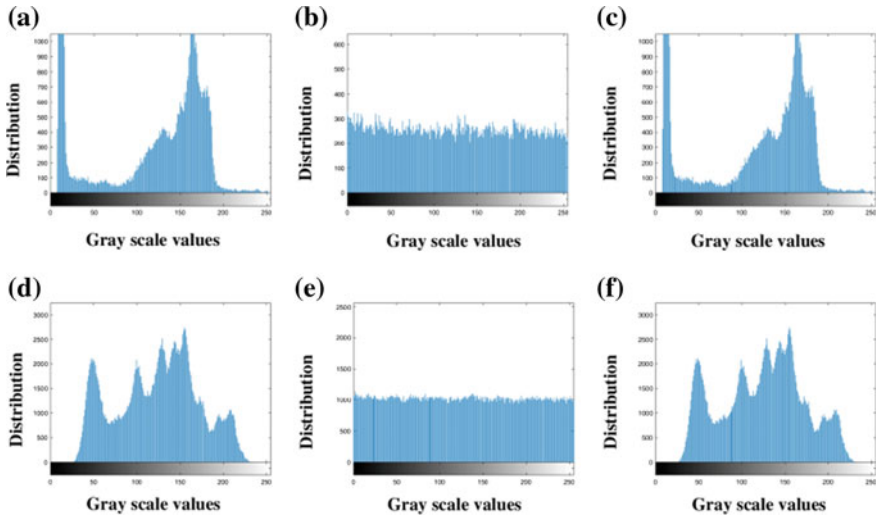
**Table 4** Results for key space

Algorithm	key space
Our algorithm	$1.963 \times 2^{426}$
Ref. [25]	More than $2^{349}$
Ref. [26]	$3.4 \times 10^{80} \approx 1.4337 \times 2^{267}$

from traces of original images. The variation between cipher image histograms (uniform distribution) and plain image histograms (nonuniform distribution) [27–29] is required. Figure 3 depicts the histogram results. By observing Fig. 3, computer output result we can say the histogram of cipher image is uniform. Hence, statistical attack is not possible in the proposed method.

#### 4.2.2 Histogram Variance Analysis

Histogram variance quantitatively analyzes the consistency of pixel values in histogram outputs. Variance and consistency of pixel values are inversely proportional. For better encryption scheme consistency is high by having low variance. The results in Table 5 prove that the innovated scheme is more protective than the schemes in [30, 31].



**Fig. 3** Simulation results for histogram: plain-cipher-decrypted images of “Cameraman” are (a), (b), and (c), respectively, and “Lena” are (d), (e), and (f)

**Table 5** Variance comparison results

Algorithm		Original image	Encrypted image
Our algorithm	Cameraman	1.1097e + 05	254.0313
	Lena	6.3340e + 05	1.0905e + 03
Ref. [30]	Lena	–	5554.8293
Ref. [31]	Lena	–	5335.8309

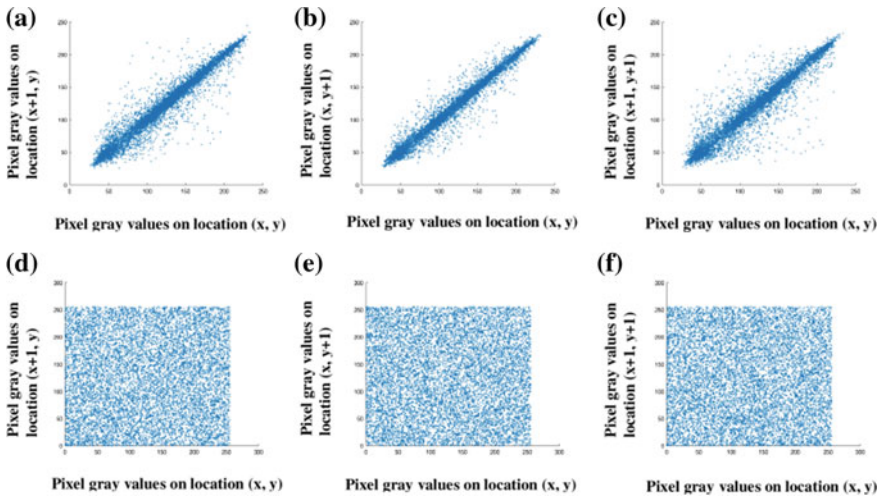
### 4.2.3 Correlation Analysis

It is a measure of the amount of interdependence among adjacent intensity values in images. For plain images correlation coefficient have to be close to +1 or −1, on other hand the value for encrypted images have to be around 0. Table 6 depicts the correlation coefficient values. We are able to achieve correlation values around zero for the encrypted “Lena” image using the suggested scheme. These values are better than values in the schemes in [32, 33].

Figure 4 is having the correlation plots for “Lena” image in horizontal, vertical, and diagonal directions. We can see from the cipher image of “Lena”, the neighboring intensity values are very weakly dependent on each other, whereas in the original image they are mostly interdependent. This provides resistant to statistical attacks.

**Table 6** Correlation coefficient comparison results

Algorithm		Our algorithm		Ref. [32]	Ref. [33]
		Camerman	Lena	Lena	Lena
Original image	Diag.	0.9034	0.9577	0.9570	0.9448
	Horz.	0.9343	0.9717	0.9597	0.9761
	Vert.	0.9572	0.9859	0.9792	0.9626
Cipher image	Diag.	0.0096	0.0014	0.0504	0.0013
	Horz.	0.0004	0.0027	0.1257	-0.0285
	Vert.	0.0019	0.0072	0.0581	0.0014



**Fig. 4** Correlation outputs in directions of **a** Horz.,—(a) and (d); **b** Vert.,—(b) and (d); **c** Diag.;—(c) and (f) for Original “Lena” and Encrypted “Lena” images, respectively

### 4.3 Differential Attack Analysis

Actually, this analysis includes two techniques. They are NPCR which stands for Number of Pixels Changing Rate and UACI which means Unified Average Changing Intensity (UACI). They analyze how much strong a scheme is against differential attacks. Better the value of NPCR/UACI, better is the algorithm. The average UACI and NPCR values for one hundred images are shown in Table 7. Altering one random pixel in plain image we get the results for above analysis. We obtained good NPCR and UACI values than the processes in [34, 35].



**Table 7** Comparison of average NPCR and UACI results

Algorithm		Average UACI (%)	Average NPCR (%)
Our scheme	Cameraman	33.4476	99.6077
	Lena	33.4593	99.6097
Ref. [34]	Lena	33.41	99.60
Ref. [35]	Lena	33.4342	99.607

**Table 8** Comparison of information entropy results

Algorithms		Plain images	Cipher images
Our algorithm	Lena	7.4451	7.9993
	Cameraman	7.0097	7.9972
Ref. [35]	Lena	–	7.9894

#### 4.4 Information Entropy Analysis

The uncertainty of gray level values in the images is calculated through this analysis. Higher is the value, higher is the changeability in pixels. The ideal value for a 256-gray image is 8. The Table 8 shows that the suggested scheme attains a value of 7.9972 and 7.9993 (almost 8 in both cases) for “Cameraman” and “Lena” images respectively has better values compared to the algorithm in [35]. This shows the innovation is strong against the entropy attacks.

#### 4.5 Known-Plaintext Attack (KPA) and Chosen-Plaintext Attack (CPA) Analysis

KPA and CPA are common cryptography attacks. In KPA the cryptanalyst is having part of plaintext and its corresponding ciphertext, then cryptanalyst utilizes this plain and ciphertext combination to decrypt the image. Similarly, is the case with CPA but in this cryptanalyst, will have access to choose the plaintext of his choice. All the keys that are used in this encryption process are dependent over plaintext. We generate them using the SHA-256 algorithm. Since the input plaintext changes the keys generated using it also changes. So, through KPA and CPA the attacker might get one key but can't use that key to decrypt the whole image because the keys we generate is highly dependent over plaintext which alter every time.

## 5 Conclusion

Through this work, we suggest a way to encrypt images using dual time DNA encoding and decoding and PWLCM systems. In this technique, three different PWLCM systems are used. The first PWLCM system is used for DNA rule selection, the second PWLCM system is used for DNA shuffling operation, and the third PWLCM system is used for key image generation. The proposed technique is quite simple and provides more confusion in the encryption process. The computer simulations and security analyses prove that suggested scheme has good encryption effect, high resistance toward statistical attack, entropy, and differential attacks, KPA and CPA attack and large secret key space. From the results, we can infer that the proposed innovation is having high security in comparison with other schemes. Since this scheme is having better results, it is more reliable and efficient for image encryption.

**Acknowledgements** We thanks Information Security Education Awareness (ISEA) project phase – II, Ministry of Electronics and Information Technology (MeitY), Govt. of India.

## References

1. Coppersmith, D. (1994). The data encryption standard (DES) and its strengths against attacks. *IBM Journal of Research and Development*, 38(3), 243–250.
2. Pub NF. 197. (2001, November 26). *Advanced encryption standard (AES)*. Federal Information Processing Standards Publication 197, US Department of Commerce/NIST.
3. Gao, H., Zhang, Y., Liang, S., & Li, D. (2006). A new chaotic algorithm for image encryption. *Chaos, Solitons & Fractals*, 29(2), 393–399.
4. Gupta, A., Thawait, R., Patro, K. A. K., & Acharya, B. (2016). A novel image encryption based on bit-shuffled improved tent map. *International Journal of Control Theory and Applications*, 9(34), 1–16.
5. Samhita, P., Prasad, P., Patro, K. A. K., & Acharya, B. (2016). A secure chaos-based image encryption and decryption using crossover and mutation operator. *International Journal of Control Theory and Applications*, 9(34), 17–28.
6. Shadangi, V., Choudhary, S. K., Patro, K. A. K., & Acharya, B. (2017). Novel Arnold scrambling based CBC-AES image encryption. *International Journal of Control Theory and Applications*, 10(15), 93–105.
7. Matthews, R. (1989). On the derivation of a chaotic encryption algorithm. *Cryptologia*, 13(1), 29–42.
8. Liu, W., Sun, K., & Zhu, C. (2016). A fast image encryption algorithm based on chaotic map. *Optics and Lasers in Engineering*, 84, 26–36.
9. Rui, L. (2015). New algorithm for color image encryption using improved 1D logistic chaotic map. *Open Cybernetics & Systemics Journal*, 9(1), 210–216.
10. Jiang, J., & Yin, Z. (2013). The advantages and disadvantages of DNA password in the contrast to the traditional cryptography and quantum cryptography. In *Proceedings of The Eighth International Conference on Bio-Inspired Computing: Theories and Applications* (pp. 307–316). Berlin and Heidelberg: Springer.
11. Zhang, J., Fang, D., Ren, H. (2014). Image encryption algorithm based on DNA encoding and chaotic maps. *Mathematical Problems in Engineering*.

12. Xue, X., Zhang, Q., Wei, X., Guo, L., & Wang, Q. (2010). An image fusion encryption algorithm based on DNA sequence and multi-chaotic maps. *Journal of Computational and Theoretical Nanoscience*, 7(2), 397–403.
13. Liu, L., Zhang, Q., & Wei, X. (2012). A RGB image encryption algorithm based on DNA encoding and chaos map. *Computers & Electrical Engineering*, 38(5), 1240–1248.
14. Jain, A., & Rajpal, N. (2016). A robust image encryption algorithm resistant to attacks using DNA and chaotic logistic maps. *Multimedia Tools and Applications*, 75(10), 5455–5472.
15. Xiang, T., Liao, X., & Wong, K. W. (2007). An improved particle swarm optimization algorithm combined with piecewise linear chaotic map. *Applied Math and Computation*, 190(2), 1637–1645.
16. Zhang, X., & Wang, X. (2017). Multiple-image encryption algorithm based on mixed image element and permutation. *Optics and Lasers in Engineering*, 92, 6–16.
17. Patro, K. A. K., & Acharya, B. (2018). Secure multi-level permutation operation based multiple colour image encryption. *Journal of Information Security and Applications*, 40, 111–133.
18. Patro, K. A. K., Acharya, B., & Nath, V. (2018). A secure multi-stage one-round bit-plane permutation operation based chaotic image encryption. *Microsystem Technologies*, 1–8.
19. Patro, K. A. K., Banerjee, A., & Acharya, B. (2017). A simple, secure and time efficient multi-way rotational permutation and diffusion based image encryption by using multiple 1-D chaotic maps. In *International Conference on Next Generation Computing Technologies* (pp. 396–418). Singapore: Springer.
20. Naskar, P. K., & Chaudhuri, A. (2016). Secured secret sharing technique based on chaotic map and DNA encoding with application on secret image. *Imaging Science Journal*, 64(8), 460–470.
21. Patro, K. A. K., & Acharya, B. (2018). Novel data encryption scheme using DNA computing. In *Advances of DNA computing in cryptography* (pp. 69–110). Chapman and Hall/CRC.
22. USC-SIPI image database for research in image processing, image analysis, and machine vision. Retrieved September 19, 2017, from <http://sipi.usc.edu/database/>.
23. Floating-point Working Group. (1985). IEEE standard for binary floating-point arithmetic. ANSI, IEEE Std. (pp. 754–1985).
24. Kulsoom, A., Xiao, D., & Abbas, S. A. (2016). An efficient and noise resistive selective image encryption scheme for gray images based on chaotic maps and DNA complementary rules. *Multimedia Tools and Applications*, 75, 1–23.
25. El-latif, A. A. A., Li, L., Zhang, T., Wang, N., Song, X., & Niu, X. (2012). Digital image encryption scheme based on multiple chaotic systems. *Sensing and Imaging*, 13, 67–88.
26. Guesmi, R., Farah, M. A. B., Kachouri, A., & Samet, M. (2016). A novel chaos-based image encryption using DNA sequence operation and secure hash algorithm SHA-2. *Nonlinear Dynamics*, 83(3), 1123–1136.
27. Wang, X., & Zhang, H. L. (2016). A novel image encryption algorithm based on genetic recombination and hyper-chaotic systems. *Nonlinear Dynamics*, 83(1–2), 333–346.
28. Mohanty, S., Shende, A., Patro, K. A. K., & Acharya, B. (2017). A DNA based chaotic image fusion encryption scheme using LEA–256 and SHA–256. *Indian Journal of Scientific Research*, 14(2), 190–201.
29. Sravanthi, D., Patro, K. A. K., Acharya, B., & Majumder, S. (2019). A secure chaotic image encryption based on bit-plane operation. In *Soft Computing in Data Analytics* (pp. 717–726). Singapore: Springer.
30. Zhu, Z., Zhang, W., Wong, K., & Yu, H. (2011). A chaos-based symmetric image encryption scheme using a bit-level permutation. *Information Sciences (Ny)*, 181(6), 1171–1186.
31. Zhang, Y.-Q., & Wang, X.-Y. (2014). A symmetric image encryption algorithm based on mixed linear–nonlinear coupled map lattice. *Information Sciences (Ny)*, 273, 329–351.
32. Huang, C. K., & Nien, H. H. (2009). Multi chaotic systems based pixel shuffle for image encryption. *Optics Communication*, 282(11), 2123–2127.
33. Chai, X. (2017). An image encryption algorithm based on bit level Brownian motion and new chaotic systems. *Multimedia Tools and Applications*, 76(1), 1159–1175.

34. Brindha, M., & Ammasai Gounden, N. (2016). A chaos based image encryption and lossless compression algorithm using hash table and Chinese remainder theorem. *Applied Soft Computing Journal*, 40, 379–390.
35. Wu, X., Kurths, J., & Kan, H. (2017). A robust and lossless DNA encryption scheme for color images. *Multimedia Tools and Applications*.

# Simple Permutation and Diffusion Operation Based Image Encryption Using Various One-Dimensional Chaotic Maps: A Comparative Analysis on Security



Dasari Sravanthi, K. Abhimanyu Kumar Patro , Bibhudendra Acharya   
and M. Prasanth Jagapathi Babu

**Abstract** With the development in technology, the security of transmission and storage of digital information (basically, digital images) is a challenge to all cryptographic researchers. In recent years, for securing digital images multiple encryption techniques have been proposed. Among them, one-dimensional (1D) chaotic map based image encryption techniques render better security in storage and transmission of images. 1D chaotic maps are simple in structure and hence efficient to implement in both software and hardware. An image encryption based on simple permutation and diffusion operation using various 1D chaotic maps is proposed in this paper. The proposed technique first performs pixel permutation operation using various 1D chaotic maps and then performs pixel diffusion operation using pixel key generated by the Secure Hash Algorithm-256 and the plain image. Also in this paper, a comparative analysis of security is presented using various 1D chaotic maps in image encryptions. The comparative results show the best security of using most of the 1D chaotic maps in image encryptions.

**Keywords** Image encryption · Permutation · Diffusion · One-dimensional chaotic maps · Secure hash algorithm SHA-256 · Security analysis

---

D. Sravanthi

Department of Information Technology, National Institute of Technology Raipur, Raipur, India  
e-mail: [sravanthi.dasari1994@gmail.com](mailto:sravanthi.dasari1994@gmail.com)

K. A. K. Patro · B. Acharya (✉) · M. Prasanth Jagapathi Babu

Department of Electronics and Telecommunication Engineering, National Institute of Technology Raipur, Raipur, India  
e-mail: [bacharya.etc@nitrr.ac.in](mailto:bacharya.etc@nitrr.ac.in)

K. A. K. Patro

e-mail: [kpatro.phd2016.etc@nitrr.ac.in](mailto:kpatro.phd2016.etc@nitrr.ac.in)

M. Prasanth Jagapathi Babu

e-mail: [jagapathi.matta@gmail.com](mailto:jagapathi.matta@gmail.com)

# 1 Introduction

Nowadays, lots of image data are being communicated through the Internet. Security of those image data is of prime importance; hence efficient encryption algorithms are required. Traditional encryption algorithms like RSA, AES, and DES [1, 2] are not appropriate to encrypt images due to their high redundancy, bulky data, strong correlation between pixels, etc. [3–6].

For the past few years, chaotic maps are being used in image encryptions to provide high security to the images. Chaotic maps have several inherent properties such as ergodicity, non-periodicity, pseudo-randomness of chaotic sequences, sensitivity to system parameters, initial conditions, etc. [7–9]. Chaotic maps are of two categories: 1D chaotic maps and high-dimensional chaotic maps [10]. The 1D chaotic maps have simple structure, easy to implement in both hardware and software, highly efficient, and have less computational cost compared to high-dimensional chaotic maps [10, 11]. So, 1D chaotic maps are preferred while implementing image encryption algorithms.

In recent years, many types of 1D chaotic maps are used for the implementation of encryption algorithms. Some of the 1D chaotic maps based image encryption techniques are as follows. In [12], Pareek et al. proposed a Logistic map based image encryption technique. In this technique, image encryption is performed using two logistic maps. In [13], Wang et al. proposed a dynamic S-box based block encryption scheme. In this technique, a Tent map is used for generating the dynamic S-boxes. In [14], Li et al. proposed a chaotic Tent map based image encryption method. In [15], Belazi et al. proposed S-box generation method based on a chaotic Sine map. The Sine map generated S-box increases confidentiality in the substitution stage of image encryption system. In [16], Zhou et al. proposed three 1D chaotic maps, namely, Logistic-Tent map, Tent-Sine map, and Logistic-Sine map in their image encryption scheme. These three new 1D chaotic maps produce larger chaotic range and better chaotic behavior as compared to their seed maps such as Sine map, Logistic map, Tent map. An encryption scheme using new Beta chaotic map was proposed by Zahmoul et al. [17]. There are a number of advantages in Beta map which are high number of system parameters, better pseudo-random chaotic sequences, strong chaotic behavior, and large range of bifurcation parameter. A new 1D chaotic map based image encryption scheme was proposed by Alpar et al. [18]. This map exhibits chaotic behavior in small interval of real numbers.

An encryption technique based on simple permutation and diffusion operation was proposed in this paper to encrypt images and various 1D chaotic maps are used in permutation operation to check the performance of each of the 1D chaotic maps. The main contributions in this paper are as follows.

- 1D chaotic maps are used to make the algorithm simpler, stronger, and more software and hardware efficient.
- Various types of 1D chaotic maps are used to check the performance of each of the chaotic maps in the proposed method.

- Simple permutation and diffusion operations are performed to add more simplicity in the algorithm.
- The SHA-256 hash algorithm is used to resist the method against known-plaintext attack (KPA) and chosen-plaintext attack (CPA).

The remaining parts of the paper are as follows. Section 2 briefly explains the one-dimensional chaotic maps used in this paper. Section 3 illustrates the simple method of the proposed encryption algorithm. Section 4 illustrates the computer simulation and security analysis results of the algorithm using various 1D chaotic maps and also in this section, the comparison analysis on security using various 1D chaotic maps are presented. In Sect. 5, the conclusion of the paper is presented.

## 2 One-Dimensional Chaotic Maps

To analyze the security of the encryption method, various 1D chaotic maps using permutation operations are performed in this scheme. The 1D chaotic maps are Kent map (KM) [19, 20], Logistic-Sine map (LSM) [16], Logistic map (LM) [21, 22], Improved Logistic map (ILM) [11], Logistic-Tent map (LTM) [16], Tent map (TM) [23], Tent-Sine map (TSM) [16], Sine map (SM) [24], Cosinus-Arcsinus Map (CAM) [25], Sinus-power Logistic map (SPLM) [25, 26], Dyadic map (DM) [27, 28], Gauss iterated map (GIM) [29], Alpar's map (AM) [18], Beta map (BM) [17].

## 3 Proposed Algorithm

### 3.1 Key Generation for Permutation

Process for secret key generation for permutation is as described below.

**Step 1:** Take an image ' $I_g$ ' of dimensions  $M_{I_g} \times N_{I_g}$

**Step 2:** SHA-256 hash algorithm is used to generate the 256-bits hash value by taking the plain image ' $I_g$ ' as input. Convert these hash values into 64-hex values which are denoted as

$$hx = hx_1, hx_2, \dots, hx_{63}, hx_{64} \quad (1)$$

**Step 3:** Convert above-produced hex values into decimal values by taking 2 hexadecimal digits of each. Therefore, 32-decimal values are produced and are represented as

$$hd = hd_1, hd_2, \dots, hd_{31}, hd_{32} \quad (2)$$

**Step 4:** Generate various 1D chaotic map based keys of the algorithm.

The generation of LM-based keys is

$$\begin{cases} xxl = xl + \left( \left( \frac{\text{mod}((hd_1 : hd_{12}), 256)}{2^9} \right) \times 0.1 \right) \\ rrl = rl + \left( \left( \frac{\text{mod}((hd_{13} : hd_{24}), 256)}{2^9} \right) \times 0.1 \right) \end{cases} \quad (3)$$

where ' $rl$ ' is the given (original) and ' $rrl$ ' is newly generated system parameters of LM, respectively; ' $xl$ ' is the given (original) and ' $xxl$ ' is newly generated initial values of LM, respectively.

Likewise, the keys (original initial value ' $xt$ ', original system parameter ' $rt$ ', generated initial value ' $txt$ ', and generated system parameter ' $rrt$ ') for TM, the keys (original initial value ' $xs$ ', original system parameter ' $rs$ ', generated initial value ' $xts$ ', and generated system parameter ' $rrs$ ') for SM, the keys (original initial value ' $xls$ ', original system parameter ' $rls$ ', generated initial value ' $xxls$ ', and generated system parameter ' $rrls$ ') for LSM, the keys (original initial value ' $xlt$ ', original system parameter ' $rlt$ ', generated initial value ' $xxlt$ ', and generated system parameter ' $rrlt$ ') for LTM, the keys (original initial value ' $xts$ ', original system parameter ' $rts$ ', generated initial value ' $xtts$ ', and generated system parameter ' $rrts$ ') for TSM, the keys (original initial value ' $xk$ ', original system parameter ' $rk$ ', generated initial value ' $xxk$ ', and generated system parameter ' $rrk$ ') for KM, the keys (original initial value ' $xcas$ ', original system parameter ' $rcas$ ', generated initial value ' $xxcas$ ', and generated system parameter ' $rrcas$ ') for CAM, the keys (original initial value ' $xspl$ ', original system parameter ' $rspl$ ', generated initial value ' $xxspl$ ', and generated system parameter ' $rrspl$ ') for SPLM are generated.

The generation of ILM-based keys is

$$\begin{cases} xxil = xil + \left( \left( \frac{\text{mod}((hd_1 : hd_8), 256)}{2^9} \right) \times 0.1 \right) \\ rril = ril + \left( \left( \frac{\text{mod}((hd_9 : hd_{16}), 256)}{2^9} \right) \times 0.1 \right) \\ kkil = kil + \left( \left( \frac{\text{mod}((hd_{17} : hd_{24}), 256)}{2^9} \right) \times 0.1 \right) \end{cases} \quad (4)$$

where ' $xil$ ' is the taken initial value and (' $ril$ ', ' $kil$ ') are the system parameters of ILM. ' $xxil$ ' is the generated initial value and (' $rril$ ', ' $kkil$ ') are system parameters of ILM.

Likewise, the keys (original initial value ' $xg$ ', original system parameters ' $alphag$ ' and ' $betag$ ', newly formed value ' $xxg$ ' and generated system parameter ' $ag$ ' and ' $bg$ ') for GIM, the keys (original initial value ' $xa$ ', original system parameters ' $aa$ ' and ' $bb$ ', generated initial value ' $xxa$ ', and generated system parameter ' $aaa$ ' and ' $bbb$ ') for AM are generated.



The generation of DM based key is

$$xxd = xd + \left( \left( \frac{\text{mod}((hd_1 : hd_{24}), 256)}{2^9} \right) \times 0.1 \right) \quad (5)$$

where ‘ $xd$ ’ and ‘ $xxd$ ’ represent the taken and generated initial values of DM, respectively.

The BM based keys are generated by using hash values ( $hd_1 : hd_2$ ) for the initial value (‘ $xb$ ’ → original, ‘ $xxb$ ’ → generated) and hash values ( $hd_3 : hd_4$ ), ( $hd_5 : hd_6$ ), ( $hd_7 : hd_9$ ), ( $hd_{10} : hd_{12}$ ), ( $hd_{13} : hd_{15}$ ), ( $hd_{16} : hd_{18}$ ), ( $hd_{19} : hd_{21}$ ), ( $hd_{22} : hd_{24}$ ) for the system parameters (‘ $x1b$ ’ → original, ‘ $xx1b$ ’ → generated), (‘ $x2b$ ’ → original, ‘ $xx2b$ ’ → generated), (‘ $b1b$ ’ → original, ‘ $bb1b$ ’ → generated), (‘ $c1b$ ’ → original, ‘ $cc1b$ ’ → generated), (‘ $b2b$ ’ → original, ‘ $bb2b$ ’ → generated), (‘ $c2b$ ’ → original, ‘ $cc2b$ ’ → generated), (‘ $kb$ ’ → original, ‘ $kbb$ ’ → generated), (‘ $ab$ ’ → original, ‘ $aab$ ’ → generated), respectively.

### 3.2 Key Generation for Diffusion

The key generation operation for diffusion is as follows.

**Step 1:** Take the last 8 hash decimal values  $hd_{25}, hd_{26}, \dots, hd_{31}, hd_{32}$  to generate the diffusion key.

**Step 2:** Apply mod-256 operation on the hash decimal values. Let  $t1-t8$  are the outputs of the mod operation.

**Step 3:** Convert  $t1-t8$  into binary. Let the binary values be denoted as  $v_1, v_2, \dots, v_7, v_8$ .

**Step 4:** Collect the first bit of each binary values  $v_1, v_2, \dots, v_7, v_8$  to generate a decimal value  $k$ .  $k$  is the key for diffusion operation.

### 3.3 Encryption Technique

Figure 1 illustrates the schematic diagram of the proposed encryption method.

**Step 1:** Take an image ‘ $Ig$ ’ of size  $M_{Ig} \times N_{Ig}$  as input.

**Step 2:** Generate permutation keys of the algorithm.

**Step 3:** Iterate the chaotic map  $M_{Ig} \times N_{Ig}$  times. The iterated chaotic sequence of LM is denoted as ‘ $XXL$ ’. The ‘ $XXL$ ’ is

$$XXL = (xxl(1), xxl(2), \dots, xxl(M_{Ig} \times N_{Ig})) \quad (6)$$

Similarly, the iterated chaotic sequences of TM, SM, LSM, LTM, TSM, ILM, KM, CAM, SPLM, DM, GIM, AM, and BM are denoted as ‘ $XXT$ ’, ‘ $XXS$ ’, ‘ $XXLS$ ’, ‘ $XXLT$ ’, ‘ $XXTS$ ’, ‘ $XXIL$ ’, ‘ $XXK$ ’, ‘ $XXCAS$ ’, ‘ $XXSPL$ ’, ‘ $XXD$ ’, ‘ $XXG$ ’, ‘ $XXA$ ’, ‘ $XXB$ ’, respectively.



**Step 6:** Diffuse the pixels of the shuffled image ‘ $P$ ’ using the diffusion key ‘ $k$ ’ obtained in Sect. 3.2 by the following process.

$$\begin{cases} C(1) = k \oplus P(1) \\ C(2) = C(1) \oplus P(2) \\ C(3) = C(2) \oplus P(3) \\ \vdots \\ C(M_{Ig} \times N_{Ig}) = C(M_{Ig} \times N_{Ig} - 1) \oplus P(M_{Ig} \times N_{Ig}) \end{cases} \quad (8)$$

where  $C(1), C(2), C(3), \dots, C(M_{Ig} \times N_{Ig})$  are the diffused pixels generated by Eq. (8).

**Step 7:** Cipher image ‘ $C$ ’ is produced by combining all the diffused pixels.

By executing the above steps in reverse order with cipher image as input, we get the decrypted image that is input plain image as output.

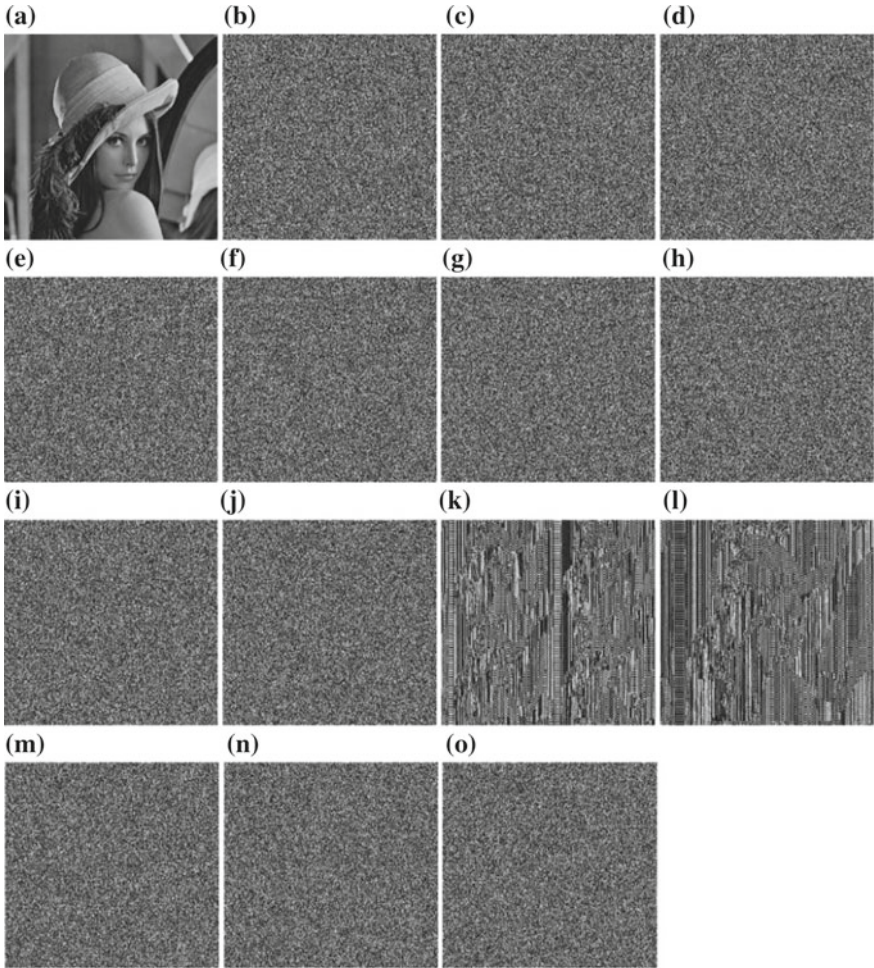
## 4 Computer Simulations and Security Analysis

This algorithm uses 14 different 1D chaotic maps to compare their performance when used in the encryption process. Computer simulations and security analysis are performed on a computer having system configurations INTEL CORE i3, 4 GB RAM, 2.40 GHz processor, Windows 7, 32-bit operating system. MATLAB R2012a is used to perform the simulation operation. This algorithm uses “Lena” image of dimensions  $512 \times 512$  to compare the performance of each of the 1D chaotic maps. Table 1 shows the initial values and system parameters of various 1D chaotic maps. From Fig. 2. We can infer that the “Lena” image is encrypted properly by using all the 1D chaotic maps except the chaotic maps—SPLM (Fig. 2k) and DM (Fig. 2l). All the images required for the simulation process are taken from image database USC-SIPI [30].

The security analyses are as given below.

**Table 1** System parameters and initial values (keys) of chaotic maps

Maps	Keys	Maps	Keys
LM	$xl = 0.2, rl = 3.9$	TM	$xt = 0.2, rt = 3.9$
SM	$xs = 0.2, rs = 3.9$	LSM	$xls = 0.2, rls = 3.9$
LTM	$xlt = 0.2, rlt = 3.9$	TSM	$xts = 0.2, rts = 3.9$
ILM	$xil = 0.2, ril = 3.9, kil = 8$	KM	$xk = 0.2, rk = 0.1$
CAM	$xcas = 0.2, rcas = 3.9$	SPLM	$xspl = 0.2, rspl = 3.4$
DM	$xd = 0.2$	AM	$xa = 2.3, aa = 0.5, bb = 2$
GIM	$xg = 0.2, \text{alphag} = 4.9, \text{betag} = -0.58$		
BM	$xb = -0.21, x1b = -0.73, x2b = 1, b1b = 8, c1b = 1, b2b = 3, c2b = -1, kb = 0.89, ab = -0.23$		



**Fig. 2** “Lena” image simulation results: **a** plain image, cipher images using **b** LM, **c** TM, **d** SM, **e** LSM, **f** LTM, **g** TSM, **h** ILM, **i** KM, **j** CAM, **k** SPLM, **l** DM, **m** GIM, **n** AM, **o** BM

#### 4.1 Key Space Analysis

Different keys utilized in this scheme are,

- The system parameters and initial values of 1D chaotic maps.
- The 256-bits hash value generated by SHA-256 hash algorithm.

According to IEEE,  $10^{-15}$  is data representation [31] standard for floating points. So, the suggested method uses a secret key with accuracy of  $10^{-15}$ . To defend against the common attacks SHA algorithm produces key space of  $2^{128}$ . In the proposed method, the total key space using various 1D chaotic maps is as shown in Table 2.

**Table 2** key space results of various 1D chaotic maps

Maps	Total key space	Result
LM	$(10^{15} \times 10^{15}) \times 2^{128} = 1.5768 \times 2^{227}$	$>2^{128}$
TM	$(10^{15} \times 10^{15}) \times 2^{128} = 1.5768 \times 2^{227}$	$>2^{128}$
SM	$(10^{15} \times 10^{15}) \times 2^{128} = 1.5768 \times 2^{227}$	$>2^{128}$
LSM	$(10^{15} \times 10^{15}) \times 2^{128} = 1.5768 \times 2^{227}$	$>2^{128}$
LTM	$(10^{15} \times 10^{15}) \times 2^{128} = 1.5768 \times 2^{227}$	$>2^{128}$
TSM	$(10^{15} \times 10^{15}) \times 2^{128} = 1.5768 \times 2^{227}$	$>2^{128}$
ILM	$(10^{15} \times 10^{15} \times 10^{15}) \times 2^{128} = 1.4000 \times 2^{277}$	$>2^{128}$
KM	$(10^{15} \times 10^{15}) \times 2^{128} = 1.5768 \times 2^{227}$	$>2^{128}$
CAM	$(10^{15} \times 10^{15}) \times 2^{128} = 1.5768 \times 2^{227}$	$>2^{128}$
SPLM	$(10^{15} \times 10^{15}) \times 2^{128} = 1.5768 \times 2^{227}$	$>2^{128}$
DM	$10^{15} \times 2^{128} = 1.7758 \times 2^{177}$	$>2^{128}$
GIM	$(10^{15} \times 10^{15} \times 10^{15}) \times 2^{128} = 1.4000 \times 2^{277}$	$>2^{128}$
AM	$(10^{15} \times 10^{15} \times 10^{15}) \times 2^{128} = 1.4000 \times 2^{277}$	$>2^{128}$
BM	$(10^{15} \times 10^{15} \times 10^{15} \times 10^{15} \times 10^{15} \times 10^{15} \times 10^{15} \times 10^{15} \times 10^{15}) \times 2^{128} = 1.3722 \times 2^{576}$	$\gg 2^{128}$

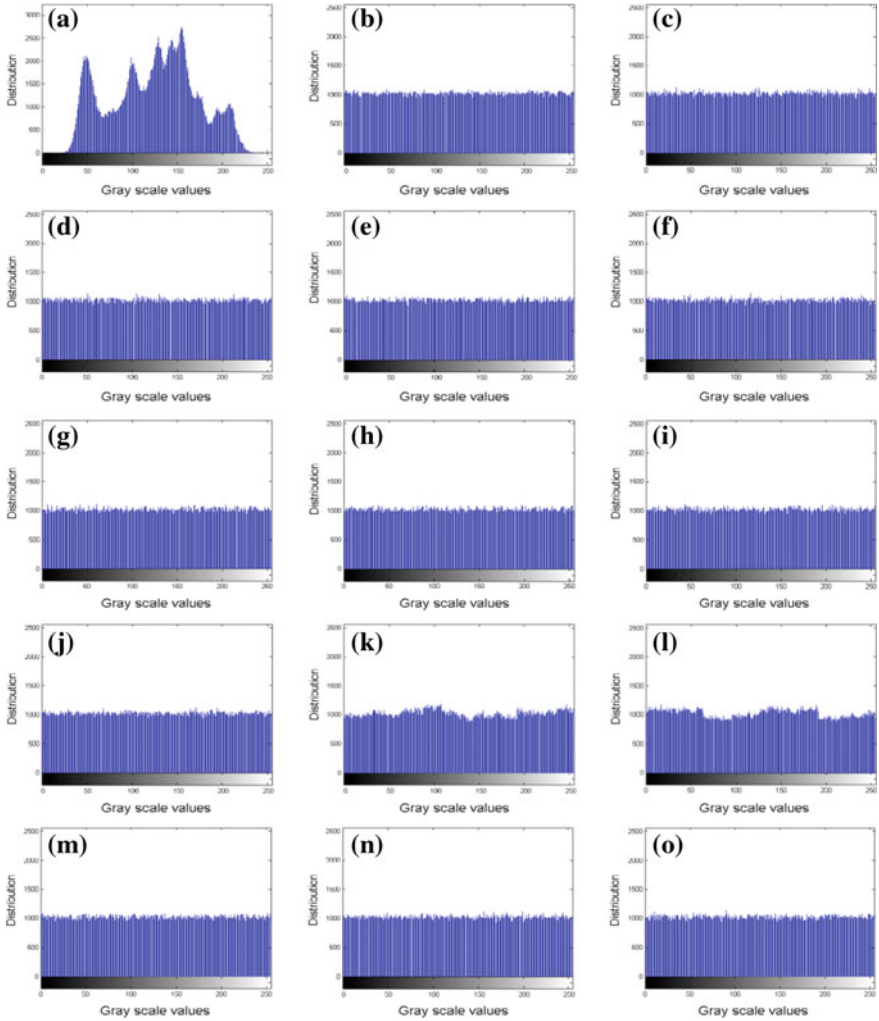
In the table, it is seen that the key space of all the chaotic maps are more than  $2^{128}$  to evade brute-force attack [32, 33]. The key space for Beta map is very larger than  $2^{128}$  which shows high resistivity of Beta map based encryption algorithm against brute-force attack.

## 4.2 Statistical Attack Analysis

**Histogram Analysis.** It is a graphical display of the distribution of data. Uniform histograms make it difficult for the attacker to extract the information of the images [34–39]. Figure 3 depicts the histogram results of “Lena” image using various 1D chaotic maps. In Fig. 3 we can also find the consistency of intensity values in the histogram, using all the 1D chaotic maps except the chaotic maps—SPLM and DM. In these maps based encryption systems, the distribution of grayscale values is nonuniform. Hence, except SPLM and DM based methods, statistical attack is not possible in the proposed method.

**Histogram Variance Analysis.** Histogram variance quantitatively analyzes the consistency of pixel values in the histogram images. Lesser the value of histogram variance corresponds to the high consistency of pixel values in the histogram [40].

Table 3 shows the variance results of “Lena” image using various 1D chaotic maps. SPLM and DM based image encryption systems are less effective as compared to



**Fig. 3** “Lena” image histogram results: **a** plain image, cipher images using **b** LM, **c** TM, **d** SM, **e** LSM, **f** LTM, **g** TSM, **h** ILM, **i** KM, **j** CAM, **k** SPLM, **l** DM, **m** GIM, **n** AM, **o** BM

**Table 3** Comparison of histogram variance results of “Lena” encrypted image

Maps	LM	TM	SM	LSM	LTM	TSM	ILM
Variance	853.7	1186.5	1068.5	946.8	1121.7	1046.5	923.3
Maps	KM	CASM	SPLM	DM	GIM	AM	BM
Variance	1072.1	951.1	3818.2	5077.9	955.5	986.3	1129.0

**Table 4** Correlation distribution comparison results of “Lena” image

Maps	Plain images (3000 pairs of pixels)			Cipher images (3000 pairs of pixels)		
	Hz.	Vt.	Dg.	Hz.	Vt.	Dg.
LM	0.9734	0.9849	0.9598	0.0205	0.0133	-0.0198
TM	0.9734	0.9849	0.9598	-0.0411	0.0005	0.0166
SM	0.9734	0.9849	0.9598	0.0233	-0.0056	-0.0087
LSM	0.9734	0.9849	0.9598	-0.0201	0.0231	-0.0277
LTM	0.9734	0.9849	0.9598	-0.0269	-0.0034	-0.0291
TSM	0.9734	0.9849	0.9598	-0.0051	0.0253	-0.0438
ILM	0.9734	0.9849	0.9598	-0.0040	0.0347	-0.0075
KM	0.9734	0.9849	0.9598	0.0051	0.0224	0.0059
CAM	0.9734	0.9849	0.9598	-0.0011	0.0154	-0.0003
SPLM	0.9734	0.9849	0.9598	-0.0019	0.0466	0.0059
DM	0.9734	0.9849	0.9598	0.0328	0.0015	-0.0459
GIM	0.9734	0.9849	0.9598	0.0344	0.0088	-0.0229
AM	0.9734	0.9849	0.9598	-0.0442	0.0223	0.0085
BM	0.9734	0.9849	0.9598	-0.0108	0.0274	0.0172

other 1D chaotic map based image encryption systems since they are having high variance values when compared to other maps.

**Correlation Analysis.** This is a measure of correlations among adjacent pixels. For better encryption, it should be close to +1 or -1 and 0 for plain images and encrypted images respectively. The correlation values shown in Table 4 for the plain “Lena” image are close to +1 while in an encrypted “Lena” image, the coefficient is close to 0 along horizontal (Hz.), vertical (Vt.) and diagonal (Dg.) directions. This shows that all 1D chaotic maps based proposed encryption system resist statistical attack.

### 4.3 Differential Attack Analysis

Actually, this analysis includes two techniques. They are Number of Pixel Changing Rate (NPCR) and Unified Average Changing Intensity (UACI). Table 5 shows the results of the two techniques. The minimum, maximum, and average values for 100 encrypted images are shown in Table 5, where these values are computed by altering any one grayscale value in the plain image. This analysis outputs are very close or more than the ideal NPCR (99.6094%) and UACI (33.4635%). The results show that the suggested scheme efficiently resists differential attack using all the 1D chaotic maps.

**Table 5** UACI and NPCR comparison results of “Lena” image

Maps	NPCR in %			UACI in %		
	Minimum	Maximum	Average	Minimum	Maximum	Average
LM	99.5916	99.6248	99.6104	33.4104	33.5263	33.4637
TM	99.5992	99.6246	99.6110	33.4097	33.5064	33.4640
SM	99.5997	99.6258	99.6112	33.4111	33.5075	33.4578
LSM	99.5968	99.6173	99.6074	33.4572	33.5217	33.4829
LTM	99.5955	99.6253	99.6125	33.4073	33.4905	33.4654
TSM	99.5944	99.6193	99.6077	33.4186	33.4748	33.4440
ILM	99.5912	99.6278	99.6088	33.4182	33.5303	33.4630
KM	99.5950	99.6287	99.6111	33.4330	33.5420	33.4745
CAM	99.5948	99.6220	99.6091	33.4169	33.5274	33.4654
SPLM	99.5626	99.6497	99.6095	32.6680	35.6438	33.4972
DM	99.0070	100.0000	99.7186	7.1829	50.5084	34.6092
GIM	99.5944	99.6241	99.6090	33.3883	33.5170	33.4574
AM	99.5936	99.6279	99.6085	33.3918	33.5934	33.4681
BM	99.5911	99.6208	99.6090	33.3905	33.5479	33.4658

### 4.4 Information Entropy Analysis

For an ideal scheme, the information entropy value is near to 8, more it is nearer to 8, more is the randomness. Table 6 is the results of the information entropy of “Lena” image which is encrypted using various 1D Chaotic maps has a value nearer 8. This shows that all the 1D chaotic maps based encryption system provide a greater degree of randomness of pixels in images. We can also infer that the SPLM and DM based encryption systems provide lesser degree of randomness in pixels compared to the other 1D chaotic maps based encryption systems.

**Table 6** Information entropy comparison results of “Lena” image

Maps	LM	TM	SM	LSM	LTM	TSM	ILM
Entropy	7.9994	7.9992	7.9993	7.9993	7.9992	7.9993	7.9993
Maps	KM	CASM	SPLM	DM	GIM	AM	BM
Entropy	7.9994	7.9993	7.9974	7.9965	7.9993	7.9993	7.9992



#### 4.5 *Known-Plaintext Attack (KPA) and Chosen-Plaintext Attack (CPA)*

In Known-plaintext attack, the attacker has some plain text and its corresponding ciphertext. He extracts the properties of the algorithm by using them. In Chosen-plaintext attack, the attacker encrypt some random plaintext by using algorithm and try to decode keys. In this proposed method, all the keys used are dependent on the input plaintext. Therefore keys will be changing with plaintext. So it is not possible to decode keys and it is resistant to KPA and CPA.

#### 4.6 *Comparison Analysis*

In this analysis, the proposed encryption system using various 1D chaotic maps are compared based on various attacks. Table 7 shows the comparative analysis results. The comparison is between simulation outputs (SO), information entropy analysis (IEA), brute-force attack (BFA), differential analysis (DA), Histogram and variance analysis (HVA), correlation analysis (CA), CPA and KPA. In Table 7, we can see that histogram attack may feasible in the SPLM and DM based encryption systems. In addition to this, in SPLM and DM based encryption systems, the image is not properly encrypted to get better encryption results.

**Table 7** Comparison of different attacks using various 1-d chaotic maps

Maps	SO	BFA	H VA	CA	DA	IEA	KPA & CPA
LM	✗	✗	✗	✗	✗	✗	✗
TM	✗	✗	✗	✗	✗	✗	✗
SM	✗	✗	✗	✗	✗	✗	✗
LSM	✗	✗	✗	✗	✗	✗	✗
LTM	✗	✗	✗	✗	✗	✗	✗
TSM	✗	✗	✗	✗	✗	✗	✗
ILM	✗	✗	✗	✗	✗	✗	✗
KM	✗	✗	✗	✗	✗	✗	✗
CAM	✗	✗	✗	✗	✗	✗	✗
SPLM	✓	✗	✓	✗	✗	✗	✗
DM	✓	✗	✓	✗	✗	✗	✗
GIM	✗	✗	✗	✗	✗	✗	✗
AM	✗	✗	✗	✗	✗	✗	✗
BM	✗	✗	✗	✗	✗	✗	✗

✗ → Attack may not feasible, ✓ → Attack may feasible

## 5 Conclusion

A simple permutation and diffusion operation based image encryption algorithm for evaluating performances of various 1D chaotic maps is described here. This image encryption algorithm firstly used pixel permutation operation using various 1D chaotic maps and then performed pixel diffusion operation to obtain a ciphered image. Present work is a comparison among various 1D chaotic maps based on various security analyses. The SPLM and DM based encryption systems provide less security in terms of simulation results and histogram attack as compared to other 1D chaotic maps. Except SPLM and DM chaotic map, we can use the other referred maps in near future to generate stronger multiple-image encryption techniques

**Acknowledgements** We thank the Information Security Education Awareness (ISEA) Project Phase—II, Ministry of Electronics and Information Technology (MeitY), Govt. of India.

## References

1. Coppersmith, D. (1994). The data encryption standard (DES) and its strengths against attacks. *IBM Journal of Research and Development*, 38(3), 243–250.
2. Pub NF. 197. (2001, November 26). *Advanced Encryption Standard (AES)*. Federal Information Processing Standards Publication 197, US Department of Commerce/NIST.
3. Gao, H., Zhang, Y., Liang, S., & Li, D. (2006). A new chaotic algorithm for image encryption. *Chaos, Solitons & Fractals*, 29(2), 393–399.
4. Gupta, A., Thawait, R., Patro, K. A. K., & Acharya, B. (2016). A novel image encryption based on bit-shuffled improved tent map. *International Journal of Control Theory and Applications*, 9(34), 1–16.
5. Samhita, P., Prasad, P., Patro, K. A. K., & Acharya, B. (2016). A secure chaos-based image encryption and decryption using crossover and mutation operator. *International Journal of Control Theory and Applications*, 9(34), 17–28.
6. Shadangi, V., Choudhary, S. K., Patro, K. A. K., & Acharya, B. (2017). Novel Arnold scrambling based CBC-AES image encryption. *International Journal of Control Theory and Applications*, 10(15), 93–105.
7. Chai, X. (2017). An image encryption algorithm based on bit level Brownian motion and new chaotic systems. *Multimedia Tools and Applications*, 76(1), 1159–1175.
8. Guesmi, R., Amine, M., Farah, B., et al. (2016). Hash key-based image encryption using crossover operator and chaos. *Multimedia Tools and Applications*, 75(8), 4753–4769.
9. Guesmi, R., Farah, M. A. B., Kachouri, A., & Samet, M. (2016). A novel chaos-based image encryption using DNA sequence operation and secure hash algorithm SHA-2. *Nonlinear Dynamics*, 83(3), 1123–1136.
10. Liu, W., Sun, K., & Zhu, C. (2016). A fast image encryption algorithm based on chaotic map. *Optics and Lasers in Engineering*, 84, 26–36.
11. Rui, L. (2015). New algorithm for color image encryption using improved 1D logistic chaotic map. *Open Cybernetics & Systemics Journal*, 9(1), 210–216.
12. Pareek, N. K., Patidar, V., & Sud, K. K. (2006). Image encryption using chaotic logistic map. *Image and Vision Computing*, 24(9), 926–934.
13. Wang, Y., Wong, K. W., Liao, X., & Xiang, T. (2009). A block cipher with dynamic S-boxes based on tent map. *Communications in Nonlinear Science and Numerical Simulation*, 14(7), 3089–3099.

14. Li, C., Luo, G., Qin, K., & Li, C. (2017). An image encryption scheme based on chaotic tent map. *Nonlinear Dynamics*, 87(1), 127–133.
15. Belazi, A., & El-Latif, A. A. A. (2017). A simple yet efficient S-box method based on chaotic sine map. *Optik—International Journal for Light and Electron Optics*, 130, 1438–1444.
16. Zhou, Y., Bao, L., & Chen, C. P. (2014). A new 1D chaotic system for image encryption. *Signal Processing*, 97, 172–182.
17. Zahmoul, R., Ejbali, R., & Zaied, M. (2017). Image encryption based on new Beta chaotic maps. *Optics and Lasers in Engineering*, 96, 39–49.
18. Alpar, O. (2014). Analysis of a new simple one dimensional chaotic map. *Nonlinear Dynamics*, 78(2), 771–778.
19. Feldman, D. P. (2012). *Chaos and fractals: An elementary introduction*. Oxford University Press.
20. Wang, X., & Wang, Q. (2014). A novel image encryption algorithm based on dynamic S-boxes constructed by chaos. *Nonlinear Dynamics*, 75(3), 567–576.
21. May, R. M. (1976). Simple mathematical models with very complicated dynamics. *Nature*, 261(5560), 459.
22. MathWorld—A Wolfram Web Resource Homepage. Retrieved May 4, 2018, from <http://mathworld.wolfram.com/LogisticEquation.html>.
23. CRAMPIN, M., & Heal, B. (1994). On the chaotic behaviour of the tent map. *Teaching Mathematics and its Applications: An International Journal of the IMA*, 13(2), 83–89.
24. Griffin, J. (2013). *The sine map*. Retrieved May 4, 2018, from <https://people.maths.bris.ac.uk/~macpd/ads/sine.pdf>.
25. Mollaeefar, M., Sharif, A., & Nazari, M. (2017). A novel encryption scheme for colored image based on high level chaotic maps. *Multimedia Tools and Applications*, 76(1), 607–629.
26. Zhang, X., & Cao, Y. (2014). A novel chaotic map and an improved chaos-based image encryption scheme. *The Scientific World Journal*, 2014.
27. Driebe, D. J. (2013). *Fully chaotic maps and broken time symmetry* (Vol. 4). Springer Science & Business Media.
28. VEPŠTAS, L. *The gauss-kuzmin-wirsing operator*. Retrieved May 4, 2018, from <http://www.linias.org/math/gkw.pdf>.
29. Hilborn, R. C. (2004). *Chaos and nonlinear dynamics: An introduction for scientists and engineers* (2nd ed.). New York: Oxford University Press.
30. USC-SIPI image database for research in image processing, image analysis, and machine vision. Retrieved September 19, 2017, from <http://sipi.usc.edu/database/>.
31. Floating-point Working Group. (1985). *IEEE standard for binary floating-point arithmetic*. ANSI, IEEE Std. (pp. 754–1985).
32. Kulsoom, A., Xiao, D., & Abbas, S. A. (2016). An efficient and noise resistive selective image encryption scheme for gray images based on chaotic maps and DNA complementary rules. *Multimedia Tools and Applications*, 75(1), 1–23.
33. Patro, K. A. K., & Acharya, B. (2018). Secure multi-level permutation operation based multiple colour image encryption. *Journal of Information Security and Applications*, 40, 111–133.
34. Wang, X., & Zhang, H. L. (2016). A novel image encryption algorithm based on genetic recombination and hyper-chaotic systems. *Nonlinear Dynamics*, 83(1–2), 333–346.
35. Patro, K. A. K., Acharya, B., & Nath, V. (2018). A secure multi-stage one-round bit-plane permutation operation based chaotic image encryption. *Microsystem Technologies*, 1–8.
36. Patro, K. A. K., & Acharya, B. (2018). Novel data encryption scheme using DNA computing. In *Advances of DNA computing in cryptography* (pp. 69–110). Chapman and Hall/CRC.
37. Sravanthi, D., Patro, K. A. K., Acharya, B., & Majumder, S. (2019). A secure chaotic image encryption based on bit-plane operation. In *Soft computing in data analytics* (pp. 717–726). Singapore: Springer.
38. Patro, K. A. K., Banerjee, A., & Acharya, B. (2017). A simple, secure and time efficient multi-way rotational permutation and diffusion based image encryption by using multiple 1-D chaotic maps. In *International Conference on Next Generation Computing Technologies* (pp. 396–418). Singapore: Springer.

39. Mohanty, S., Shende, A., Patro, K. A. K., & Acharya, B. (2017). A DNA based chaotic Image fusion encryption scheme using LEA-256 and SHA-256. *Indian Journal of Scientific Research*, 14(2), 190-201.
40. Chai, X., Chen, Y., & Broyde, L. (2017). A novel chaos-based image encryption algorithm using DNA sequence operations. *Optics and Lasers in Engineering*, 88, 197-213.

# Integration of Wireless Sensor Networks with Cloud Towards Efficient Management in IoT: A Review



Rajendra Kumar Dwivedi, Nikita Kumari and Rakesh Kumar

**Abstract** Internet-of-things (IoT) became very popular in today's research. IoT means all devices of a particular system should be connected with each other through the internet. Cloud Computing and Wireless Sensor Networks (WSN) are integrated for efficient management in IoT. This integration is known as Sensor Cloud. This technology has a lot of applications due to the continuous development of information and communication technology. Although sensor cloud has several advantages still it has many research challenges like energy efficiency, security, QoS, etc. The wireless sensor network is the network of sensors which operate on battery. Reducing energy consumption and communication overhead are important issues of wireless sensor networks. Efficient management of WSN and cloud results in efficient management of IoT. This paper presents a survey on efficient management of IoT with sensor cloud.

**Keywords** IoT · Cloud computing · WSN · Sensor cloud · Virtualization

## 1 Introduction

IoT, cloud computing, and WSN are the latest technologies which can optimize an application with help of each other. Integration of WSN with cloud is called sensor cloud [1–3]. Figure 1 shows the architecture of sensor cloud. At lowest layer, there are physical sensors which are mapped with virtual sensors at middle layer with help of cloud. The upper layer consists of end-users who can run multiple applications at a time with same WSN. End users can also use more than one WSN at a time

---

R. K. Dwivedi · N. Kumari (✉) · R. Kumar  
Department of Computer Science and Engineering, Madan Mohan Malaviya University of  
Technology, Gorakhpur, UP, India  
e-mail: [niki06790@gmail.com](mailto:niki06790@gmail.com)

R. K. Dwivedi  
e-mail: [rajendra.gkp@gmail.com](mailto:rajendra.gkp@gmail.com)

R. Kumar  
e-mail: [rkiitr@gmail.com](mailto:rkiitr@gmail.com)

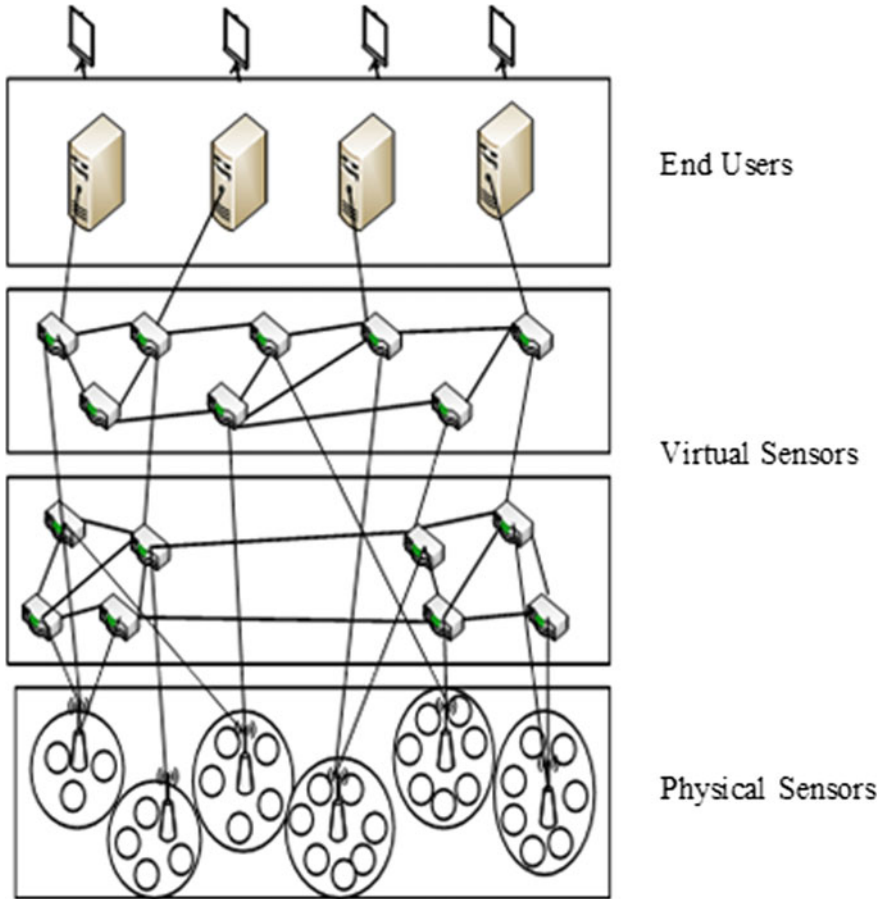


Fig. 1 Architecture of sensor cloud

within the same application with help of virtualization managed by cloud. Thus, cloud can provide sensor-as a-service. This integration helps not only the cloud but also helps WSNs [4–6] to store and effectively manage the sensor data at cloud. Other advantages of cloud include low cost of maintenance, flexibility of services, fault-tolerant communication, backup, recovery, etc. Cloud computing also enables on-demand sensor networks that can be released with minimal management efforts. Several IoT applications are based on sensors. Therefore, this integration can also help IoT for its efficient management. This integration helps a lot in real-time applications. It has several advantages and opportunities still it has some issues and challenges too, such as security [7–10], energy efficiency [11–14], load management [15], etc. Nowadays, many types of research are being carried out in this field. This paper presents a review on such researches toward efficient management of IoT.

Rest of the paper is organized as follows. Section 2 describes some basic terminologies. Various applications of IoT are discussed in Sect. 3. Section 4 focuses on related work. Finally, Sect. 5 concludes the paper with some research directions.

## **2 Basic Terminologies**

Few preliminaries and basic terminologies are discussed below:

### ***2.1 Wireless Sensor Network***

Wireless sensor networks are popular because of their capability of building their own network for several environment monitoring as well as military applications [16]. They have a small size, processing capability, and memory [17, 18]. WSNs are created to sense various physical phenomenon like light, temperature, humidity, radiation, sound, etc. Wireless sensor network helps to provide a bridge between the physical and virtual world. It has a very large range of applications in industries, transportations, infrastructures, military, etc. Wireless sensor network can be explained as a self-configured framework. Its various applications monitor physical or environmental conditions to collect the sensed data. This network also has several constraints such as power, memory, processing capability, etc. Global positioning system and local positioning algorithms are used to get location and position information.

### ***2.2 Cloud Computing***

Cloud computing is the delivery of computing services such as server, storage, software, platform, database, networking, etc. It is based on pay-per-use technique [15]. This computing method provides various on-demand computer services available over the internet. Cloud computing is approaching to experience direct cost and it is expected to transform a data center from a capital-intensive set up to a variant price environment. Cloud computing modifies the equivalent traditional concepts of grid computing and distributed computing. It is a pool of abstracted, extremely scalable and control computing infrastructure capability of host-end client request and billed by a managed process [6, 19, 20].

### **2.3 *Internet of Things***

Most of the applications of IoT are based on sensors which monitor the physical and environmental conditions [1, 21, 22]. A cloud also helps to store and process bulk of data generated by an IoT application. In IoT at remote a command is given which controls the capabilities of the device. IoT devices are often mobile and can be deployed at various locations. They need to be connected to server side from a lot of different places. Internet-of-things is a network of physical devices which are based on internet. The internet is not only a network of computers it has spread into a network of device of all types and sizes such as smartphone, medical instrument, industrial system, etc. All such devices are connected, communicate, and share information based on some protocols in order to obtain smart reorganizations, positioning, tracing, safety, and control.

## **3 IoT Applications**

IoT is the need of today's life. There can be many IoT-enabled applications such as smart parking, smart animal farming, smart waste management system, etc. Some of the applications of this technology are shown in Fig. 2 and discussed below:

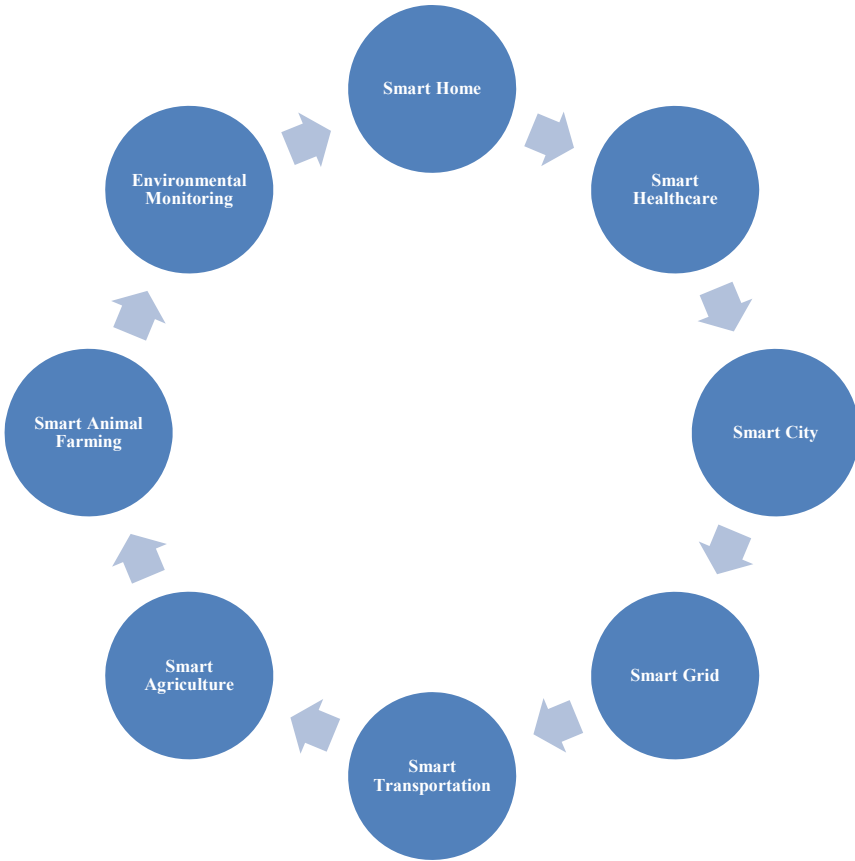
### **3.1 *Smart Home***

Our life-style at home can be improved by making it smart. A smart home facilitates us in many ways. A smart house can automatically down the blinds of window or close the window as per the requirements of the seasons.

### **3.2 *Smart Health Care***

Healthcare application is gaining popularity these days where actuators and embedded sensors are implanted in the patients to receive the health related data of the patients. This data is analyzed automatically and doctors can provide facilities and health related services to the patients accordingly [4]. Thus, suitable actions can be taken for the betterment of the patients by the healthcare system.





**Fig. 2** Major applications of IoT

### **3.3 Smart City**

Good quality of life in the smart city is improved by providing comfort and ease to the residents. Interest for information is received according to people’s necessary different interconnect systems that understand and offer the suitable services (like transport, utilizes health, etc.) to people.

### **3.4 Smart Grid**

Smart grid is techniques that provides electricity from supplier to consumer through digital technology for saving energy and reduces cost and increase reliability. The network operator is all about the extension of grid observation, improved reliability,

wide measurement, and self-healing properties. The system integrated it all about integration of it and automation application.

### ***3.5 Smart Transportation***

The smart transportation systems are generally developed by the government or transportation authority. This system controls the traffic with help of smart traffic signal system. Various progressive programs regulate transportation. The advancement of electric vehicles, charging facility and dedicated short range communication helps the development of smart vehicular system.

### ***3.6 Smart Agriculture***

Smart agriculture is a computing-based agriculture that helps to change and reorient the agriculture systems. It efficiently supports the development and guarantees of food security during and ever-changing climate. The main focus of smart agriculture is to enhance agriculture productivity and incomes.

### ***3.7 Smart Animal Farming***

Smart animal farming became very popular these days. It provides proper diet care and environment required for animals. Smart farming is required to monitor animal farms remotely. This intelligent system should also do surveillance of the entire farm. Good care of animals is very important.

### ***3.8 Environmental Monitoring***

Use of cloud and IoT can change the development of high-speed information system between the entities. Sensors are deployed in any monitoring area to sense some environmental conditions such as temperature, pressure, movement of objects, etc. [5]. Some monitoring can be continuous and long term such as level of water, gas concentration in air, soil humidity and other characteristics, inclination for static structures, position changes, lighting condition of infrared radiation for fire and animal detection. A cloud-based data access is able to structure the potential energy requirements of low energy communicative segments and presents fast access to the data for end user. It accepts to manage and process the complex events given by the real-time data flow of sensors.

**Table 1** IoT area, utilization, requirements and challenges

IoT area	Utilization and requirements	Challenges
Smart home	Industrial consolidation, Development of multi-power saving and cross-application	The core component, security, private protection
Smart healthcare	Medicine, treatment of remote virtual, the sharing and management and information of patient for treatment and drug	The industrial clear are no planning, limited manufactured abilities sensor, medical and biomedical largest scalability of data
Smart city	Efficient delivery of public utilities such as water, electricity as well as associated government service	It requires smart people
Smart grid	The power generated in sensor monitoring, the power supply in automatically management	The core lack of technology, communication including reliable, electromagnetic security, and capabilities
Smart transportation	The RFID technology in development of intelligent transport system	The transportation management from various administration department
Smart agriculture	Real-time access and information sharing of agricultural resources, intelligent management of products circulation and safety	Lack of low-cost sensing technology and devices, lack of communication infrastructure in countryside
Smart animal farming	This smart system will operate remotely for monitoring the animal farms	It will detect any misshaping and protection or such type of like fire
Environmental monitoring	Environment monitoring includes population, impressive sciences, geographic research, monitoring of flood and fire	Less number for monitoring station and least develop management platform, less in developed on manufacturing the high accuracy sensor chips, the unified industry standard

Table 1 shows the comparative study of various areas of IoT, their utilization, requirements, and challenges.

## 4 Related Work

In this section, a survey on various techniques for efficient management of IoT with sensor cloud is presented. The analysis is discussed as follows.

Madden et al. [23] described the industrial vision correctly. Sensors have limited constraints such as intermittent connectivity, energy and memory constraints etc.

Maintaining record of historical information is difficult for sensor data streams. These limitations show that the traditional database instrumentation is unsuitable for queries over sensor. They presented the Fjords architecture for query management over sensor data streams and limiting the resource demands. This architecture also helps to maintain the high query throughput.

Gnawali et al. [24] discussed CTP report, which is a variable rating protocol from the wireless sensor networks. CTP usually uses three techniques to give effective, robust, and reliability routing for high-equilibrium network condition. CTP's link estimator has accurate platform-independent interface. Second, CTP usually uses the algorithm to time to manage traffic, sending few visual signals in stability topology yet quickly adaptive to changes. Finally, CTP activates the technology with data traffic quickly to discover and fix routing failures.

Sudarshan et al. [25] demonstrated that the mobile sensor is a rising technology which is being researched in large in the past decade. This research survey paper studies the concept of mobile sensor integrated with the cloud service. It informs the different mobile sensor availability and their classification. It studies the necessary and limited mobile sensor network in terms of store computed power efficiency and scalable.

Estrin Deborah et al. [26] explained that sensed data can be stored at cloud so that it can be retrieved by any handheld devices like mobile phones anytime and from anywhere to analyze the system. Any user can fetch the data using his mobile phone or computer as per the permissions granted to him by the IoT system. Thus it helps to many expectations of daily lives. Present data capture leveraged data processing and personal data overleaps are the essential components for these emerging systems.

Alessio et al. [27] described that IoT has now become part of our life. They explained that cloud and IoT are merged together to serve a varied number of application scenarios.

Nair [28] explained that WSNs are used broadly in different areas. They discussed a model for power-aware scheme.

Dash et al. [29] presented wide range of critical applications that get and process data of remote sensor systems from the real world.

Dinh et al. [30] told that a volatile increase of the mobile application and environment communication in mobile computing is the cause of the evolution of sensor cloud.

Dash et al. [31] explained that there is an expanding pattern of utilizing distributed computing circumstance for the capacity of information process. Cloud computing gives applications, platforms, and foundation over the internet. It is another mechanism to get the shared assets. Remote sensors have been viewed as the most fundamental innovation for the twenty-first century which is spatially distributed in the sensor network for information transmission. Secure and easygoing access of information in distributed computing is very expansive.

A comparative study of the literature survey of related work is shown in Table 2.

**Table 2** Comparative study

Authors	Year	Contribution	Remarks
Madden et al. [23]	2016	Explained the industrial vision correctly	Architecture from the decision multiple queries over many sensors and should be limited sensor resources demand get through maintaining high query throughout
Gnawali et al. [24]	2009	CTP reports a variable rating protocol from the wireless sensor networks	The technology with data traffic is quick
Sudarshan et al. [25]	2010	Survey of Mobile sensor is a rising technology	Mobile sensor power efficiency and scalable
Deborah et al. [26]	2016	Mobile phone and cloud service collective and analyse system systematically data	Present data capture leveraged data processing and personal data overlap are the essential components for these emerging system
Alessio et al. [27]	2016	Technology that is both already part of our life	Where cloud and IoT are merged together is predicted as disruptively
Nair et al. [28]	2011	Environment monitoring surveillance and military applications	Management circumstance management, data conglomerations, connect management
Dash et al. [29]	2012	Distributed resources sharing	Real-time traffic accumulation, real-time environment data monitoring
Dinh et al. [30]	2013	Integrated cloud computing into the mobile environment	Performs environment communication in mobile computing
Dash et al. [31]	2010	Utilizing distributed computing circumstance	Expanding pattern of utilizing distributed computing

## 5 Conclusions and Future Directions

Today, IoT has become part of human life. Many IoT-enabled devices are available in market and still, in future, there is a large scope for researches related to IoT. Several IoT applications are based on sensors and cloud. Therefore, for efficient management of IoT, sensors, and cloud should also be efficiently managed. This paper provides a review on efficient management of IoT using sensor cloud which is the result of integration of WSNs and cloud.

There are several other research issues and challenges to work upon such as security, QoS, cost control, pricing, etc. These challenges provide future directions to the research.

## References

1. Dwivedi, R. K., Singh, S., & Kumar, R. (2019). Integration of wireless sensor networks with cloud: A review. In *2019 9th IEEE International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 115–120). Noida, India.
2. Dwivedi, R. K., & Kumar, R. (2018). Sensor cloud: Integrating wireless sensor networks with cloud computing. In *2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)* (pp. 820–825). Gorakhpur, India.
3. Dwivedi, R. K., Saran, M., & Kumar, R. (2019). A survey on security over sensor-cloud. In *2019 9th IEEE International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 31–37). Noida, India.
4. O'Brien, M. (2008). Remote telemonitoring—A preliminary review of current evidence. *European Center for Connected Health* (pp. 76–82).
5. Lazarescu, M. (2013). *Design of a WSN platform for long-term environmental monitoring for IoT applications* (pp. 45–54). IEEE Journal: Emerging and Selected Topics in Circuits and Systems.
6. Ponnagal, R. S., & Raja, J. (2011). An extensible cloud architecture model for heterogeneous sensor services. *International Journal of Computer Science and Information Security*, 9(1), 227–233.
7. Dwivedi, R. K., Sharma, P. & Kumar, R. (2018). A scheme for detection of high transmission power based wormhole attack in WSN. In *2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)* (pp. 826–831). Gorakhpur, India.
8. Dwivedi, R. K., Sharma, P. & Kumar, R. (2018). Detection and prevention analysis of wormhole attack in wireless sensor network. In *2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 727–732). Noida, India.
9. Dwivedi, R. K., Pandey, S., & Kumar, R. (2018). A study on machine learning approaches for outlier detection in wireless sensor network. In *2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 189–192). Noida, India.
10. Sharma, P., & Dwivedi, R. K. (2019). Detection of high transmission power based wormhole attack using received signal strength indicator (RSSI). In S. Verma, R. Tomar, B. Chaurasia, V. Singh, & J. Abawajy (Eds.), *Communication, networks and computing. CNC 2018. Communications in computer and information science* (Vol. 839, pp. 142–152). Singapore: Springer.
11. Kumar, P., Kumar, R., Kumar, S., & Dwivedi, R. K. (2010, November). Improved modified reverse AODV protocol. *International Journal of Computer Applications—IJCA*, 12(4), 22–26.
12. Dwivedi, R. K., Tiwari, R., Rani, D., & Shadab, S. (2012). Modified reliable energy aware routing protocol for wireless sensor network. *International Journal of Computer Science & Engineering Technology—IJCSET*, 3(4), 114–118.
13. Verma, K., & Dwivedi, R. K. (2016). A review on energy efficient protocols in wireless sensor networks. *International Journal of Current Engineering and Scientific Research—IJCESR*, 3(12), 28–34.
14. Verma, K., & Dwivedi, R. K. (2017). AREDDP: Advance reliable and efficient data dissemination protocol in wireless sensor networks. In *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)* (pp. 1–4). Coimbatore.
15. Dwivedi, R. K. (2012). From grid computing to cloud computing & security issues in cloud computing. *Technia: International Journal of Computing Science and Communication Technologies—IJCSCT*, 5(1), 805–809.
16. Naik, A. K., & Dwivedi, R. K. (2016). A review on use of data mining methods in wireless sensor network. *International Journal of Current Engineering and Scientific Research—IJCESR* (ISSN PRINT: 2393–8374, ISSN ONLINE: 2394-0697), 3(12), 13–20.
17. Agarwal, A., Maddhesiya, S., Singh, P., & Dwivedi, R. K. (2012). A long endurance policy (LEP): An improved swap aware garbage collection for NAND flash memory used as a swap space in electronic devices. *International Journal of Scientific and Engineering Research—IJSER*, 3(6), 412–417.

18. Chaudhary, M. K., Kumar, M., Rai, M., & Dwivedi, R. K. (2011, January). A modified algorithm for buffer cache management. *International Journal of Computer Applications—IJCA*, 12(12), 47–49.
19. European Commission. (2013). *Definition of a research and innovation policy lever aging cloud computing and IoT combination*. Tender specifications, SMART 2013/0037.
20. Jeffery, K. (2014). Keynote: CLOUDs: A large virtualisation of small things. In *The 2nd International Conference on Future Internet of Things and Cloud (FiCloud-2014)*.
21. He, W., Yan, G., Xu, L. D. (2014). Developing vehicular data cloud services in the IoT environment. *IEEE Transactions on Industrial Informatics*, 10(2), 1587–1595.
22. Lee, K., Murray, D., Hughes, D., & Joosen, W. (2010). Extending sensor networks into the cloud using Amazons web services. In *IEEE International Conference on Networked Embedded Systems for Enterprise Applications (NESEA)* (pp. 1–7).
23. Madden, S. R., & Franklin, M. J. (2016). Fjording the stream: An architecture for queries over streaming sensor data. In *The 18th International Conference on Data Engineering* (pp. 106–112).
24. Gnawali, O., Fonseca, R., Jamieson, K., Moss, D., & Levis, D. (2009). Collection tree protocol. In *The 7th ACM Conference on Embedded Networked Sensor Systems (SenSys)* (pp. 89–95).
25. Sudarshan, K. S. (2010). A comprehensive study of mobile sensing and cloud services. In *IEEE Conference* (pp. 117–123).
26. Deborah, E. (2016). Participatory sensing: Applications and architecture [internet predictions]. In *IEEE conference on Internet Computing* (pp. 12–42).
27. Alessio, B. (2016). Integration of cloud computing and internet of things: A survey. *Future Generation Computer Systems*, 56, 684–700.
28. Nair, G. N., Morrow, P. J. & Parr, G. (2011). Design considerations for a self-managed wireless sensor cloud for emergency response scenario (pp. 189–195).
29. Dash, K. S. (2012). Sensor-cloud: Assimilation of wireless sensor network and the cloud. In *International Conference on Computer Science and Information Technology* (pp. 193–199). Springer.
30. Dinh, H. T. (2013). A survey of mobile cloud computing: Architecture, applications, and approaches. *Wireless Communications and Mobile Computing*, 1587–1611.
31. Dash, K. S., Mohapatra, S., & Pattnaik, P. K. (2010). A survey on applications of wireless sensor network using cloud computing. *International Journal of Computer Science & Emerging Technologies*, 50–55.

# Reliability-Based Resource Scheduling Approach Using Hybrid PSO-GA in Mobile Computational Grid



Krishan Veer Singh and Zahid Raza

**Abstract** The inclusion of smartphone/mobile nodes as a part of grid computing increases the computation limits of the static grid while at the same time adds to the complexity owing to the associated factors like mobility, limited power, and weak wireless connectivity. This work presents a hybrid PSO-GA (Particle Swarm Optimization—Genetic Algorithm) based resource allocation strategy for reliable execution of jobs within a reasonable time for the computational mobile grid. Before allocating the task to the resources, the best nodes as per the fitness function are selected under the given constraints in order to meet the scheduling objectives. PSO-GA is a hybrid approach, proving to be more efficient and effective than single PSO or GA. Simulation study supports the effectiveness of the proposed approach.

**Keywords** Mobile grid computing · Reliability · Particle swarm optimization (PSO) · Genetic algorithm (GA)

## 1 Introduction

Advancement in technology cumulatively has helped the researchers to create better handheld devices with enormous computing capacity at cheaper price. The statistics shown in Fig. 1 presents research done by Statista—The Statistics Portal to observe the growth of smartphone users over years supporting the above claims [1]. It can be observed from the figure that today we have around 2.53 billion smartphone users that can contribute to research if a potential platform is created. This was the motivation for the development of the concept of Mobile Computational Grid (MCG) that can combine the huge computing power of these handheld devices across the globe with desktop grid or on the standalone basis of mobile grid to contribute to the progress of both science and society.

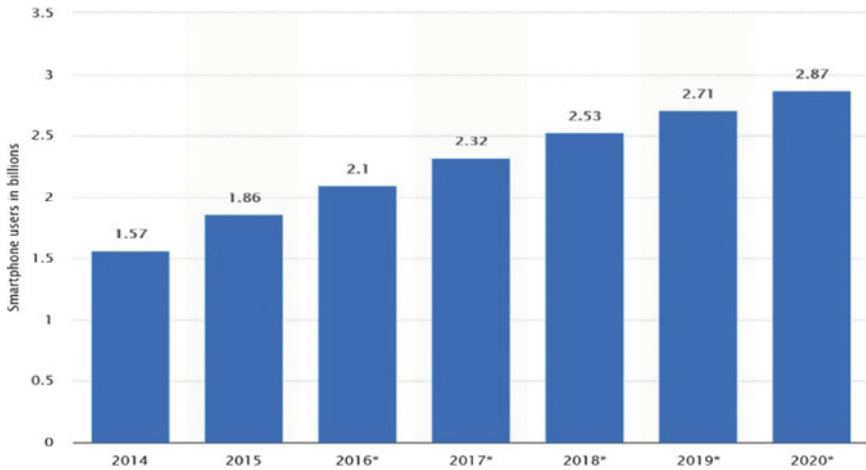
---

K. V. Singh (✉) · Z. Raza  
School of Computer and Systems Sciences, Jawaharlal Nehru University,  
New Delhi 110067, India  
e-mail: [kv.jnu07@gmail.com](mailto:kv.jnu07@gmail.com)

Z. Raza  
e-mail: [zahidraza@mail.jnu.ac.in](mailto:zahidraza@mail.jnu.ac.in)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_11](https://doi.org/10.1007/978-981-15-0694-9_11)





**Fig. 1** Worldwide number of smartphone users from 2014 to 2020 (in Billions) [4]

The mobile grid system follows all the concepts of existing grid systems [2–5]. The penetration and utility of these handheld devices coupled with the advent of the IoT paradigm promise to be enormous in future thus calling for the proper exploitation of the underutilized capacity of this whole mobile computational workforce to be used as a grid [6, 7]. MCG includes the grid infrastructure when computing nodes are mainly mobile nodes but even existing grid of desktop computers can be included to contribute toward a more effective system. The application of mobile grid includes the research projects to even day-to-day applications, e.g., translating book into any other languages, music, medical services, data mining, creating a big picture for better rescue and help operation of a calamity affected [8, 9].

Being an NP class of problem, various Quality of Service (QoS) parameters have been used. This encouraged us in this work to look into the problem considering the reliable job allocation for the MCG while considering the other domain constraints. Accordingly, this work proposes a hybrid PSOGA method to schedule the tasks on the available MCG resources by keeping in mind maximization of the job execution reliability in the system.

## 2 Scheduling Model

The work is based on certain assumptions for better real-life simulation scenario. We consider the following assumptions to simulate the proposed approaches.

- Sufficient bandwidth is available and no network congestion when a node moves to another cell.
- Nodes participating in grid infrastructure have a better battery life cycle and have a constant decay rate.

- Nodes with a minimum of 1 GB RAM are considered for smooth processing of the request.
- Mobile node entering from another cell is not eligible for allocation if their mobility score, user behavior, and battery status are unknown.
- Nodes moving at a higher speed are not considered for allocation due to frequent cell change and a handshake, which increases its power consumption rate and hence decreases the system reliability.
- Smartphone battery stays at least for a day even if it participates in grid infrastructure and amidst mild user behavior.
- One job module can be assigned to only one node. However, the task can be distributed among various nodes distributed as modules.

## 2.1 Reliability Expression

We follow the approach proposed by [10–12] to calculate the node probability during the task execution and hence compute the reliability of the grid system. In order for a task T assigned to the nodes in the MCG to execute, it must be operational during its execution and each path must be functional during the inter-module communication (IMC) between the two processing nodes. The reliability of a processor  $P_k$  during the time interval t follows Poisson distribution and is calculated as shown below:

$$R_k(X) = e^{-\int_0^t \lambda_k(t) dt} \quad (1)$$

The expression reduces to  $e^{-\lambda_k t}$ , if we assume that the failure rate  $\lambda_k$  is constant during the operation. The total execution time for the module to  $p_k$  can be written as

$$\sum_{i=1}^n x_{ik} e_{ik} \quad (2)$$

Thus, the reliability  $R_k(T, X)$  for the execution of the task to the processors  $P_k$  can be estimated as

$$R_x(T, X) = \exp\left(-\lambda_k \sum_{i=1}^n x_{ik} e_{ik}\right) \quad (3)$$

where

- m: number of modules in the task,
- T: Tasks assigned to the processor
- $\lambda_k$ : The failure rate of processor
- X: Matrix represents an assigned task to a node
- $e_{ik}$ : Execution time of task  $t_i$  on processor  $p_k$
- $x_{ik}$ : an element of X,

$$x = \begin{cases} 1 & \text{if task } t_i \text{ is assigned to no } dep_k \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

Equation 2 gives the total time required to execute the task  $t_i$  on processor  $p_k$ . Let,  $L_{kb}$  be the path between the two processors  $p_k$  and  $p_b$ . The reliability of the communication link between these two processors during  $t$  time interval can be calculated as

$$R_{kb}(X) = e^{-\mu_{kb}t} \quad (5)$$

where  $t$  is the total time elapsed in transmitting the IMC and is evaluated as

$$R_{kb}(T, X) = \exp\left(-\mu_{kb} \sum_{i=1}^n \sum_{j=1}^n x_{ik}x_{jb} \left(\frac{c_{ij}}{w_{kb}}\right)\right) \quad (6)$$

where

$\mu_{kb}$ : the failure rate of the path  $L_{kb}$ .

$c_{ij}$ : the IMC between the task  $t_i$  and  $t_j$ ,  $c_{ij} = c_{ji}$ , and  $c_{ii} = 0$ ,  $\forall 1 \leq i, j \leq n$  and

$w_{kb}$ : the transmission rate of path  $L_{kb}$ .

The system is reliable when both its involved components are fully operational during the execution. Accordingly, the system reliability with the task allocation  $X$ , as the total system reliability can be calculated as

$$R(X) = \prod_{k=1}^n R_k(X) \prod_{k=1}^n \prod_{b>k}^{n-1} R_{kb}(X) \quad (7)$$

$$= \exp(-COST(X)) \quad (8)$$

### 3 Fitness Function

In Eq. (8),  $COST(X)$  function is the required fitness function that needs to be optimized. A negative sign indicates that in order to maximize the reliability  $R(X)$ , we need to minimize the  $COST(X)$  function expanded from Eq. (8) as shown in Eq. (9). The first component in the  $COST(X)$  represents the unreliability caused by the executing nodes of various reliabilities and the second term gives the unreliability caused by the possible link failures during the IPC between the two processors calculated using Eqs. 3 and 6. The  $COST(X)$  can then be written as

$$COST(X) = \sum_{k=1}^n \sum_{i=1}^m \lambda_k x_{ik} e_{ik} + \sum_{k=1}^{n-1} \sum_{b>k}^m \sum_{i=1}^m \sum_{j=1}^m \mu_{kb} x_{ik} x_{jb} \left(\frac{c_{ij}}{w_{kb}}\right) \quad (9)$$

## 4 The Proposed Model

The infrastructure of MCG consists of  $N$  number of mobile nodes including the likes of smartphone, laptops, tablet, etc., with processing capacity  $p_k$  ( $1 \leq k \leq N$ ) and connected through possibly weak wireless network. All these devices are dynamic in nature due to their mobility or due to the connectivity issues. Mobile devices like smartphone are mobile in nature assumed moving with velocity  $v$  and powered by a limited power source with a battery status denoted as BP. All these features and constraints have some benefits and restrictions at the same time. These issues if not properly taken care of leads to the failure of either computing nodes or the entire mobile grid system. To address this, we propose a reliability-based resource scheduling model in which  $n$  number of mobile nodes are selected to schedule the task comprising of modules. The resources are selected considering the following points:

- First, the resources are categorized on the basis of the available battery power in clusters of good, moderate, and poor resources.
- Depending on the usage pattern and mobility of nodes we divide the available nodes into these three above-mentioned categories [10].
- Considering the above conditions, we select the good and moderate resources (nodes) among available, leaving the poor resources so as to avoid the possible failure of job execution which otherwise would lead to the delay of results or no result at all.
- PSO-GA is then applied to look for the best mapping of these tasks on the mobile grid on the selected resources as mentioned above. The process terminates with the allocation of the tasks on the selected cluster of resources as suggested by PSO-GA in order to optimize the QoS which in this case is reliability.

### 4.1 Hybrid PSO-GA

This work proposes the use of hybrid PSO-GA. The hybrid approach begins from the initialization step where the population or swarm of particles and their corresponding velocities are randomly generated over the search space. However, the initial position is randomly chosen from the number of available processors  $x_0^i \in [1, n]$ , i.e., the values lie in the range ( $x_{min} = 1$  and  $x_{max} = n$ ) representing the corresponding lower and upper boundary values, respectively, where  $n$  is the total no. of usable processors under the given constraints. The velocity vector is used to update the current position of each particle in the swarm from the memory gained from each particle known as particle best ( $p_{best}$ ) and knowledge gained from the whole swarm (known as global best ( $g_{best}$ )). Calculation of velocity is shown as Eq. (10) in which the first part is the momentum that improves the ability of global search and second part corresponding to the cognitive influence that helps in learning from individual experiences.

$$v_i(t + 1) = w \cdot v_i(t) + c_1 * r_1 * (p_{best} - x_i(t)) + c_2 * r_2 * (g_{best} - x_i(t)) \quad (10)$$

The last part is known as the social influence which is the process of learning from the experiences of others representing the information sharing and social cooperation between particles. Here,  $c_1$ ,  $c_2$  are acceleration constants and  $r_1$ ,  $r_2$  are random number between (0, 1). The pseudocode of the proposed hybrid algorithm is shown in the above section [13–15].

```

PSOGA
{
  Initialize the parameters and randomly generate initial particle/population
  Set the Initial velocity of particles to zero.
  Initialize constants  $c_1$ ,  $c_2$ , ( $P_c$ ), ( $M_c$ ) and iteration count.
  Repeat the following until maximum iteration count
  Begin PSO
  Calculate the fitness value of each particle using equation (4.10).
  If fitness value > pBest
    pBest = fitness value
  end
  Choose the particle with best fitness value as gBest i.e. gBest =
  best(pBest).
  For each particle update velocity of particle and particle position
  end
end PSO

Start GA
Calculate the fitness value of each particle in the population using equation (9)
Do
  Select a pair of parent chromosomes for crossover with probability  $P_c$ .
  While (until same size of population generated)
    Select the predefined percentage of chromosomes among all to perform the
    mutation operator with probability  $M_c$ .
    Replace the parent chromosomes with the mutated chromosomes.
    Replace the old population with the new population for new iteration.
  end GA
end
Calculate the fitness value again using equation (9).
Return the best particle.
}

```

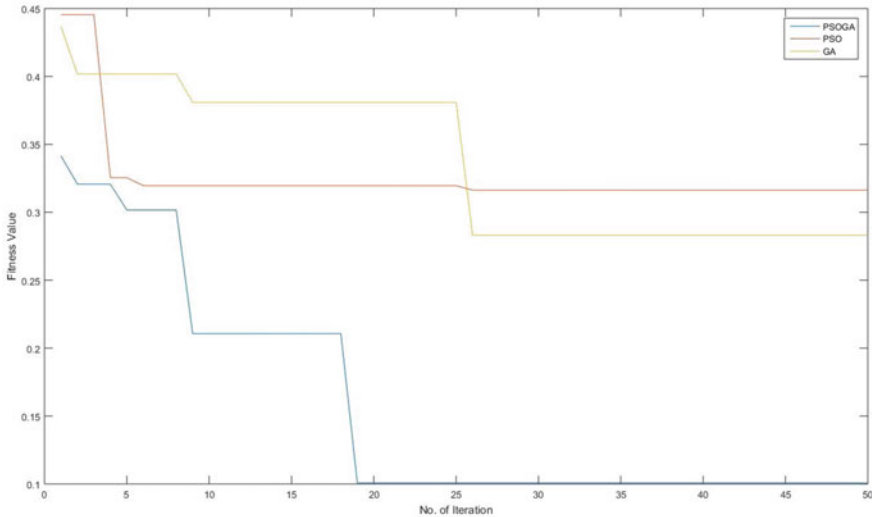
### 5 Experimental Study

This section presents the experimental study of the proposed model and the observations through the simulation study performed using MATLAB R2010b simulator. Simulation experiments use random values to generate population and other parameters but within the certain predefined range. The initial set of population is randomly generated. The values and the range of the parameters used are as shown in Table 1.

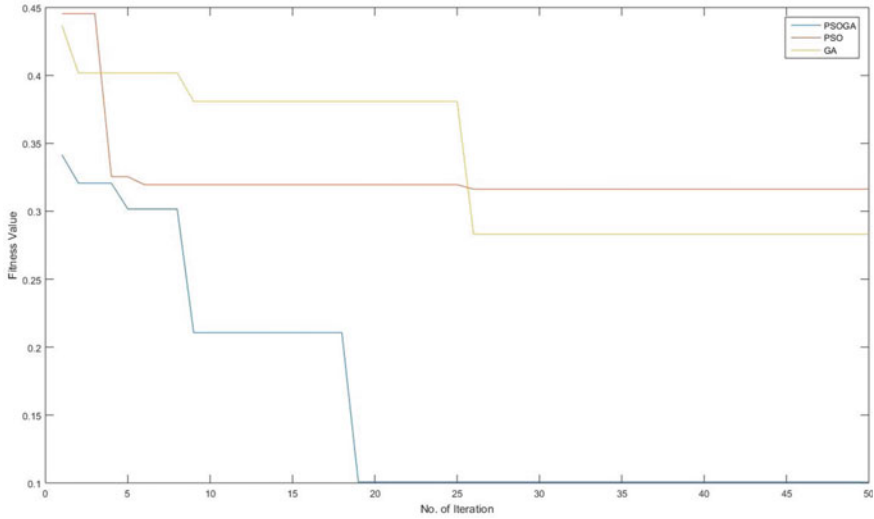
The experimental study shows that the hybrid algorithm outperforms the PSO and GA under various conditions. The experiments were run five times and the average of the results have been presented. The experiments were performed with varying Job, Nodes and Iteration counts and the results are presented as Figs. 2, 3 and 4. By varying the job size and changing the processing nodes in each case, the hybrid algorithm reports a better performance. As can be seen in all the results, the hybrid approach

**Table 1** Parameters used

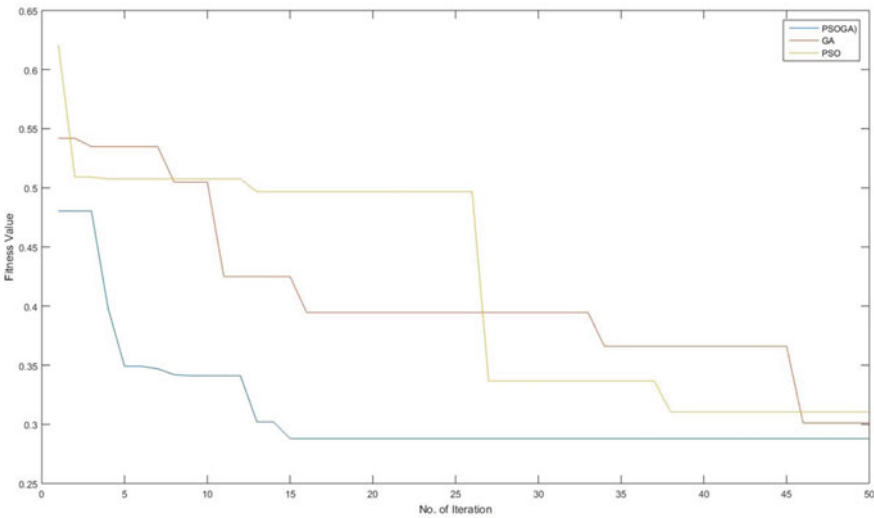
No. of processing nodes	50–100
Job Size	40–100
Lambda	0.00005–0.0010
Mu	0.00015–0.00030
Population size	30, 50, 100
Generation count	50–100
User behavior	0–40, 40–60, 60–100
Battery status	0–40, 40–80, 80–100



**Fig. 2** Fitness variations with Job Size = 30, Processing Nodes = 50 and Iteration Count = 100



**Fig. 3** Fitness variations with Job Size = 50, Processing Nodes = 100 and Iteration Count = 100



**Fig. 4** Fitness variations with Job Size = 100, Processing Nodes = 100 and Population Size = 50

using PSO-GA eventually results in a saturation value of the Fitness in terms of reliability to be better valued than that of PSO or GA alone. The saturation in all the approaches was observed to be attained till iteration count of 100. For the reduced population size of 50 as reported in Fig. 4, the results were observed to converge faster in close to 50 generations. Further, the model also observed to perform even better by varying the processing nodes and increasing the job size. The same pattern

of results was observed for other data sets as well. From the experimental study it is also observed that the proposed PSOGA approach saturates faster than GA or PSO used alone as suggested in the literature too for similar hybrid models.

## 6 Conclusion

This work studies the reliability-based job allocation on the MCG. To study the reliability, the work considered the battery power and usage pattern of the node being two crucial factors in deciding the suitability of the nodes considered for allocation and thus deciding the reliability of the system. A hybrid evolutionary approach combining PSO and GA is used to schedule the tasks on the selected mobile grid resources to attain the allocation pattern resulting in maximum reliability of the job execution. The system performance is evaluated through simulation study with the results being compared with basic PSO and Genetic Algorithm. Results indicate that the proposed hybrid approach outperforms the single evolutionary approach of PSO or GA when used alone.

## References

1. Statista—The Statistics Portal. Retrieved April 15, 2018, from <https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/>.
2. Rajaraman, V., & Murthy, C. R. (2003). *Parallel computers*. New Delhi: Architecture and Programming Prentice-Hall of India.
3. Berman, F., Fox, G., & Hey, T. (2003). *Grid computing: Making the global infrastructure a reality*. Wiley and Sons.
4. Foster, I., & Kesselman, C. (1998). *The grid—Blueprint for a new computing infrastructure*. Morgan Kaufmann.
5. Foster, I., & Kesselman, C. (2002). What is the grid? -a three-point checklist. GRID today2002.
6. Li-juan, D. U., Zhen-wei, Y. U. (2010). Development of mobile grid. *Computer Engineering and Design*, 31(6), 1166–1169.
7. Duan, L., Kubo, T., Sugiyama, K., Huang, J., Hasegawa, T., & Walrand, J. (2014). Motivating smartphone collaboration in data acquisition and distributed computing. *IEEE Transaction on Mobile Computing*, 13(10).
8. Zeng, W.-Y., Zhao, Y.-L., Zeng, J.-W., et al. (2008). Mobile grid architecture design and application. In *4th International Conference on Wireless Communications Networking and Mobile Computing*. WiCOM '08, pp. 1–4, Oct. 2008.
9. Viswanathan, H., Lee, E. K., & Pompili, D. (2012). Mobile grid computing for data- and patient-centric ubiquitous healthcare. The 1st IEEE Workshop Enabling Technologies for Smartphone and Internet Things (ETSIoT), pp. 36–41.
10. Singh, K. V., & Raza, Z. (2017). A quantum-inspired binary gravitational search algorithm-based job-scheduling model for mobile computational grid. *Concurrency and Computation: Practice and Experience*, 29(12).
11. Shatz, S. M., Wang, J. P., & Goto, M. (1992). Task allocation for maximizing reliability of distributed computer systems. *IEEE Transactions on Computers*, 41, 1156–1168.
12. Kartik, S., & Murthy, C. (1997). Task allocation algorithms for maximizing reliability of distributed computing systems. *IEEE Transfer Computer Systems*, 46, 719–724.



13. Attiya, G., & Hamam, Y. (2006). Task allocation for maximizing reliability of distributed systems: A simulated annealing approach. *Journal of Parallel and Distributed Computing*, 66, 1259–1266.
14. Kao, Y. T., & Zahara, E. (2008). A hybrid genetic algorithm and particle swarm optimization for multimodal functions. *Applied Soft Computing*, 8(2), 849–857.
15. Oshin, A. C. (2018). Hybrid PSACGA algorithm for job scheduling to minimize makespan in heterogeneous grids. In S. Bhattacharyya, S. Sen, M. Dutta, P. Biswas, H. Chattopadhyay (eds.) *Industry Interactive Innovations in Science, Engineering and Technology. Lecture Notes in Networks and Systems*, vol. 11. Springer, Singapore, 2018.

# Internet of Things (IoT) Enabling Technologies, Requirements, and Security Challenges



Shadab Alam, Shams Tabrez Siddiqui, Ausaf Ahmad, Riaz Ahmad and Mohammed Shuaib

**Abstract** Internet of Things (IoT) is an emerging technique for connecting heterogeneous technologies related to our daily needs that can affect our lives tremendously. Many architectures and applications have been proposed and implemented using IoT platform from a simple supply chain to complex life support systems. There are many obvious benefits of such networks but these systems can cause great danger to finance and life if compromised. Such issues are hindering the mass adaptation of IoT. This requires a strong architecture that can provide strong user authentication, access control as well as privacy and trust to the users of the system. The IoT network is a heterogeneous network connecting many small hardware constraint devices and also where traditional security architectures and techniques cannot be applied. Therefore, it requires a different set of specialized techniques and architecture to provide security to the IoT network. This paper focuses on the security requirements, current state of art as well as future directions in the field of IoT.

**Keywords** Internet of things · Security goals · IoT security challenges · IoT issues · IoT architecture · Applications

---

S. Alam (✉) · S. T. Siddiqui · M. Shuaib  
Department of Computer Science, Jazan University, Jizan, Saudi Arabia  
e-mail: [s4shadab@gmail.com](mailto:s4shadab@gmail.com)

S. T. Siddiqui  
e-mail: [stabrezsiddiqui@gmail.com](mailto:stabrezsiddiqui@gmail.com)

M. Shuaib  
e-mail: [talkshuaib@gmail.com](mailto:talkshuaib@gmail.com)

A. Ahmad · R. Ahmad  
Department of Computer Science, Aligarh Muslim University, Aligarh, India  
e-mail: [ausafahmad.cs@gmail.com](mailto:ausafahmad.cs@gmail.com)

R. Ahmad  
e-mail: [riaz.ahmad.tech@gmail.com](mailto:riaz.ahmad.tech@gmail.com)

## 1 Introduction

Internet of Thing (IoT) is gradually entering and affecting our lives and this trend will further increase rapidly in the coming years. IoT is not a new thing but only in the past few years academia and industry have given more attention.

There is no specific or single definition of IoT but different groups and organizations have defined differently. Things can be defined in respect of IoT as different objects having the capability to sense, interact, and communicate with each other and environment. IoT is the interconnection of these things. According to a study by Gartner, there will be 25 billion IoT devices by the year 2020 [1]. These devices or things are heterogeneous in terms of size, capability, technology, interoperability, and attributes. These devices may vary from high capacity Super Computers to devices with minimal capacity like RFID or NFC Tags.

According to another market survey report by IHS Markit, there will be 30.73 billion IoT devices that will further increase to 75.4 billion in 2025 [2]. To make a connection of these things we require special infrastructure and middleware to incorporate these devices into existing vast “Internet” network. Due to resource constraint, many traditional security techniques will not be of any help in these devices and require special attention. This paper examines the existing infrastructure available in the field of IoT, security requirements with security techniques and envisages the future requirements to make IoT safe secure for wide adaptation and security of this infrastructure.

## 2 IoT Definition

The term “Internet of Things” was coined by Kevin Ashton in 1999 for an idea to include RFID tags in supply chain management to track objects [3]. Since then, the term Internet of Things or in short generally referred to as IoT has evolved a lot and still changing so much that it has in itself become a major area of research and development. Internet of Things can be broadly defined as a network of “Things” and people which will allow connectivity among them at any time and any place where these things may be using homogeneous or entirely heterogeneous network and services [4].

“Thing” can be defined as an object which can be uniquely identified and it can be accessed anytime from anywhere [5]. The primary function of the IoT network is to uniquely identify “Things” and connect these things to the internet for collecting information, communication, and processing. Such networks are created in order to fulfill a certain objective. Such network minimizes the human intervention in fulfilling desired objectives.

### **3 IoT Enabling Technologies**

#### ***3.1 Sensor and Actuator***

A sensor is a device that detects any change or event in its environment and generates a corresponding output signal [6]. There are various types of sensors based on their applications like proximity sensor for detecting presence of any object in a specific range. Thermal sensor for sensing the temperature of an object or environment, acoustic sensor for sensing the sound, automotive sensor for sensing speed and fuel in automobile, optical sensor for sensing absence and presence of light, electric and magnetic sensor, etc [7, 8]. The sensed information is used to generate some specific output signal automatically that is used by actuators for control or movement of a specific device or system. An actuator is a device that receives a specific signal from the sensor and generates motion corresponding to these signals [9].

Sensors and actuators are the backbones of the IoT infrastructure that are used to sense information and take automatic decisions based on these inputs that help in a high level of automation.

#### ***3.2 Machine-to-Machine (M2M) Communication***

Machine-to-machine (M2M) communications is a new concept that makes automatic data transfer and communication among different machines possible. These machines can be sensors, actuators, computers and processors, and a different mobile device of varying capacity [10, 11]. Such automatic communication networks among machines are more intelligent and less dependent on the human intervention that makes it very useful in IoT applications.

#### ***3.3 Radio-Frequency Identification (RFID)***

Radio-Frequency Identification (RFID) is a wireless technology that uses electromagnetic signals for sending and detecting an object. The RFID system has three components namely; Transponder (Tag), Antenna and a transceiver (with decoder) [12].

Since its inception in 1906, RFID technology has evolved a lot and its application is very common nowadays and considered as an option to replace barcodes which can be further used to detect and uniquely identify an object for a comparatively long distance. Tags can be read-only tags that are prewritten and cannot be further modified and Write tags that have the facility to edit the information stored in it that can be uniquely identified based on the ID stored in them. Tags can be further classified as active and passive based on electromagnetic signals and battery power.

Passive tags are dependent on the electromagnetic signal generated by the reader but with limited range. Active tags have their own power and can themselves generate signals that increase their range. Although much advancement has been made in RFID technology security is still an issue in these tags because traditional encryption and other security technologies cannot be applied on such tags due to the absence or very limited power and processing ability [13]. RFID tags will be a major part of the IoT network and need to be secured for a secure IoT. Many solutions have been proposed [14] and need to be standardized to be applied in the IoT network for application in RFID security.

### ***3.4 Near-Field Communication (NFC)***

Near-field communication (NFC) technology uses the magnetic induction field and operates at 13.56 MHz radio range that enables contactless communication between two devices at a very short range of less than 10 cm [15].

The communication can be active or passive based on that if the device creates its own radio-frequency (RF) field or not, respectively. Passive devices don't have a power supply like smartcards but active devices usually have their own power source [16]. There are many security concerns like eavesdropping, data corruption, data modification, man-in-the-middle attack, etc. are common in any wireless communication protocol and NFC is also susceptible to these. Many security solutions have also suggested such issues [17] but still, there are many open questions that need to be answered before the large-scale adoption of these devices in IoT infrastructure [18].

### ***3.5 Wireless Sensor Network***

Wireless Sensor Network (WSN) is a network of wirelessly connected sensors and actors that sense the environment and perform an action in a distributed and collaborative manner. In WSN, the sensor collects information about the physical world like temperature, sound, pressure, health information, etc., and actors take appropriate decisions and perform actions based on the information collected by the sensors [19]. These sensors can be homogeneous or heterogeneous, static or dynamic, location-aware or unaware and have basic data processing abilities. The WSN should be scalable and it can be proactive or reactive based on its reaction to the environmental changes. A reactive Wireless Sensor Network can immediately react if any changes occur in the sensed environments [20]. This gives flexibility to the actors to perform activities from a remote location. Wireless Sensor Network is a basic component of the IoT network but there are many security considerations that need to be reviewed. There can be two approaches for integrating WSN into IoT network by either giving

access directly to wireless devices or using a middle layer in form of the base station or gateway [21].

### 3.6 6LoWPAN

6LoWPAN is a networking technology that combines the Internet Protocol Version 6 (IPv6) Low-power Wireless Personal Area Networks (LoWPAN). This technology enables the devices with constrained hardware resources to connect with the internet and communicate. This supports mesh topology and a very good level of reliability at the same time very low power consumption that makes it ideal for IoT applications [22].

## 4 IoT Requirements

The most important IoT requirements are: [23].

- *Interoperability*: In IoT, there will be wide-ranging devices with heterogeneous technologies involved and there should be some mechanism for these technologies and devices to operate and communicate with each other efficiently.
- *Evolvability*: There is no standard for IoT devices and network till now and not seems possible in the near future. Even if the IoT architecture is somewhat standardized then still there will scope for new technologies and devices. IoT should be in a position to evolve itself in due course of time and accommodate newer technologies.
- *Scalability*: IoT network is about 20 billion devices and it will rapidly increase with more and more standardization and security solutions in this field. The network should be scalable to any number of devices without affecting the performance.
- *Availability*: Many critical days to the functioning of human life will depend on the IoT network that requires  $24 \times 7$  availability.
- *Performance*: Generally the system have to trade-off between low power and high performance but IoT requirements are different. IoT system should give high performance at low power consumption.
- *Resiliency*: Resiliency refers to the system's ability to recover from any possible failure, resist the external threats and adapt according to the environmental changes in which it operates [24]. Providing resiliency feature in the ever-changing IoT infrastructure is very difficult and a major area of research in future.
- *Security and Privacy*: The most important requirement of any network is security and privacy. This becomes more critical due to the networking of many resource constraint devices where traditional security and encryption standards cannot be applied and need a totally different architecture to provide security and privacy to such networks.

## **5 IoT Security Challenges**

There are various security challenges and limitations related to IoT, which are affecting large-scale adaption. In this section, these challenges and limitations have been discussed in detail as well as possible solutions to overcome these challenges.

### ***5.1 Resource-Constraint Devices***

Devices connected in the IoT network are resource constraints and very little capability for storage and capability. This is a major issue for designing or applying any security framework. Standard security frameworks, cryptographic algorithms cannot be applied in such case. Specific security mechanism and framework have to be designed and applied which can be run on a very low level of hardware configuration but at the same time can sustain a high level of brute-force attacks and other cryptanalysis techniques. Middleware based frameworks are a possible solution to address such security requirements [25].

### ***5.2 Low Energy/Power Requirements***

The devices in the IoT network are resource constraints as well and they need to run on wireless mode for long service hours with limited power consumption. This restricts also to use protocols with little computational and power requirements. 6LoWPAN is one of the major techniques to support IoT requirements with low power consumption. Other solutions are low-power-embedded devices, RFID and NFC tags, but RFID and NFC tags have many security concerns that need to be addressed.

### ***5.3 Heterogeneous Devices (Interoperability)***

Various different types of devices with varying configurations and capability are connected and ad hoc networks are created that make it very difficult to devise a mechanism for interpretability. A lot of standardization works have been done by various organizations and issues but still, no global standard in place for interoperability of IoT devices and the majority of the task are de-facto or organization specific.

## **5.4 Bandwidth**

High level of connectivity and a very large number of connected devices creates a high volume of communication and data that needs high bandwidth of Internet of speed. In the majority of countries, internet speed is still very slow. New 5G technologies may address these issues and until high bandwidth data rate is not achieved, the IoT technology will remain confined to a very small area of this world. These issues will be automatically handled in the near future with the implementation of new and evolving internet technology providing cheap and fast internet connection.

## **5.5 Scalability**

The growth of IoT devices are tremendous and billions of devices are connecting in this network with passes of time. Therefore, the IoT base network should support a high level of scalability to incorporate such huge growth of IoT devices. Middleware-based architecture can be a possible solution that will separate the application layer with the physical layer implementation and number of devices connected. Such a network can be scaled by improving the middleware resource architecture.

## **5.6 Data Volumes**

Billions of connected devices will generate a very large and complex data that has to be stored and interpreted for the automation of various activities. Big Data applications have to be applied in these networks while considering the security and privacy issues. Hadoop platform is a possible major solution for IoT-based big data analytics solutions.

## **6 Conclusion**

In this paper, a broad introduction and definition of IoT have been given. This paper analyzed various enabling technologies for IoT growth and its applications. This paper further presents the various requirements and security challenges for IoT and presented various possible solutions to these challenges. Middleware-based solutions can counter major challenges that deal with the infrastructure-based solution at the same time. Hadoop platform can provide a solution for big data analytic requirements for analyzing the data. In this way, it can be inferred that IoT security challenges and issues can be easily resolved by using the existing frameworks and technologies. There are sufficient technological advancements available that if applied can provide a secure, efficient, and reliable IoT-based solutions.



## References

1. <http://www.gartner.com/newsroom/id/3165317>.
2. IoT platforms: enabling the Internet of Things March 2016 ihs.com WHITEPAPER Sam Lucero Sr. Principal Analyst, M2M and IoT.
3. Ashton, K., That 'Internet of Things' Thing. [Online]. Retrieved May 20, 2013, from <http://www.rfidjour-nal.com/articles/view?4986>.
4. Friess, P., & Guillemin, P. (2009). Internet of things strategic research roadmap. The Cluster of European Research Projects.
5. Minerva, Roberto, Biru, Abyi, & Rotondi, Domenico. (2015). *Towards a definition of the internet of things (IoT)*. Torino, Italy: IEEE Internet Initiative.
6. Rayes, A., & Salam, S. (2017). The things in IoT: Sensors and actuators. *Internet of Things From Hype to Reality*. Springer International Publishing, pp. 57–77.
7. Wikipedia, Online: <https://en.wikipedia.org/wiki/Sensor>.
8. Sensors: Online Electrical Engineering Online: <http://www.electrical4u.com/sensor-types-of-sensor/>.
9. Actuators, The Green Book. Online: <http://www.thegreenbook.com/four-types-of-actuators.htm>.
10. Chen, Min, Wan, Jiafu, & Li, Fang. (2012). Machine-to-machine communications: Architectures, standards and applications. *KSI Transaction on Internet and Information Systems*, 6(2), 480–497.
11. Wu, G., et al. (2011). M2M: From mobile to embedded internet. *IEEE Communications Magazine*, 49(4).
12. Domdouzis, Konstantinos, Kumar, Bimal, & Anumba, Chimay. (2007). Radio-frequency Identification (RFID) applications: A brief introduction. *Advanced Engineering Informatics*, 21(4), 350–355.
13. Rosenbaum, B. P. (2014). Radio frequency identification (RFID) in health care: privacy and security concerns limiting adoption. *Journal of medical systems*, 38(3), 19.
14. Peris-Lopez, P., et al. (2016). Lightweight cryptography for low-cost RFID tags. *Security in RFID and Sensor Networks*, 121–150.
15. Curran, K., Millar, A., & Mc Garvey, C. (2012). Near field communication. *International Journal of Electrical and Computer Engineering*, 2(3), 371.
16. Coskun, Vedat, Ozdenizci, Busra, & Ok, Kerem. (2013). A survey on near field communication (NFC) technology. *Wireless Personal Communications*, 71(3), 2259–2294.
17. Haselsteiner, E., & Breittfuß, K. (2006). Security in near field communication (NFC). In *Workshop on RFID security*.
18. NFC Forum: NFC and the Internet of Things, Retrieved March 05, 2017, from NFC Forum Web site. <http://nfc-forum.org/nfc-and-the-internet-of-things/>.
19. Akyildiz, Ian F., & Kasimoglu, Ismail H. (2004). Wireless sensor and actor networks: Research challenges. *Ad Hoc Networks*, 2(4), 351–367.
20. Zanjireh, M. M., & Larijani, H. (2015). A survey on centralised and distributed clustering routing algorithms for wsns. In *Vehicular Technology Conference (VTC Spring)*, 2015 IEEE 81st. IEEE.
21. Alcaraz, C. et al. (2010). Wireless sensor networks and the internet of things: Do we need a complete integration? In *1st International Workshop on the Security of the Internet of Things (SecIoT'10)*.
22. Olsson, J. (2014). 6LoWPAN demystified. *Texas Instruments*, 13.
23. Bassi, A. et al. (2013). Enabling things to talk. Designing IoT solutions with the IoT architectural reference model, 163–211.
24. Delic, K. A. (2016). On Resilience of IoT Systems: The Internet of Things (Ubiquity symposium). Ubiquity 2016, 1.
25. Batalla, J. M., Vasilakos, A., & Gajewski, M. (2017). Secure smart homes: Opportunities and challenges. *ACM Computing Surveys (CSUR)*, 50(5), 75.

# Impact of Network Load for Anomaly Detection in Software-Defined Networking



Ashish Gupta, Bharat Didwania, Gaurav Singh, Hari Prabhat Gupta, Rahul Mishra and Tanima Dutta

**Abstract** Software-Defined Networking (SDN) introduces a new network paradigm for separating the control plane and data plane. The control plane manages the packet flow in the data plane of the network. The anomaly detection in the context of SDN is to identify potentially harmful traffic. If an anomaly occurs because of malicious packets in SDN, inspecting the payload of packets is an effective way to recognize abnormal traffic. In this paper, we consider different bandwidths and topologies of the network for the detection of an anomaly in SDN. We also evaluate the performance of the SDN on the same network. We have implemented different tree topologies on OpenFlow controller using Mininet network emulator. We considered OpenFlow messages as a performance metric for evaluating the performance of the network with different tree topologies.

**Keywords** Anomaly detection · Mininet network emulator · Software-defined networking

---

A. Gupta · H. P. Gupta (✉) · R. Mishra · T. Dutta  
Department of Computer Science and Engineering, IIT (BHU), Varanasi, India  
e-mail: [hariprabhat.cse@iitbhu.ac.in](mailto:hariprabhat.cse@iitbhu.ac.in)

A. Gupta  
e-mail: [ashishg.rs.cse16@iitbhu.ac.in](mailto:ashishg.rs.cse16@iitbhu.ac.in)

R. Mishra  
e-mail: [rahulmishra.rs.cse17@iitbhu.ac.in](mailto:rahulmishra.rs.cse17@iitbhu.ac.in)

T. Dutta  
e-mail: [tanima.cse@iitbhu.ac.in](mailto:tanima.cse@iitbhu.ac.in)

B. Didwania · G. Singh  
Department of Electrical Engineering, IIT (BHU), Varanasi, India  
e-mail: [bharat.didwania.eee14@iitbhu.ac.in](mailto:bharat.didwania.eee14@iitbhu.ac.in)

G. Singh  
e-mail: [gaurav.singh.eee14@iitbhu.ac.in](mailto:gaurav.singh.eee14@iitbhu.ac.in)

# 1 Introduction

Software-Defined Networking (SDN) is a mechanism of improving network efficiency, management, and security by separating the control and data plane of the network [10, 11]. In the traditional networking concept, two important functions are as follows: first decide how to process incoming packets, e.g., to which physical port the packets will forward, and then output the data to the predecided port. In SDN, these two mentioned tasks are decoupled into data plane functions and control plane functions. The separation of control plane and data plane makes a vertical decentralization of the traditional network. The forwarding decision (control plane) is made by a controller, which manages and operates a network through open interfaces. The controller located above the data plane maintains a centralized view of the entire network, which provides a decisive advantage over the current network architectures. The logically centralized architecture supports programmability of the control plane that allows bifurcation of control plane functionality from network devices like routers, switches, etc., to specified controller instances running in a software as shown in Fig. 1.

The controllers in SDN have granular control over the switches to handle data and the ability to automatically prioritize or block certain types of packets. This increases network efficiency without investing in expensive and application-specific network switches. Multiple types of network technologies are designed for SDN that can make the network more agile and flexible to support the visualization of the server. It also supports storage infrastructure of the data center. In short, SDN can be defined as an approach for designing, building, and managing networks that creates a decentralization by separating the control plane from the data plane of the network. In other words, it is the separation of the network infrastructure and the control functions of the network.

**Fig. 1** Illustration of the SDN Architecture with anomaly detection

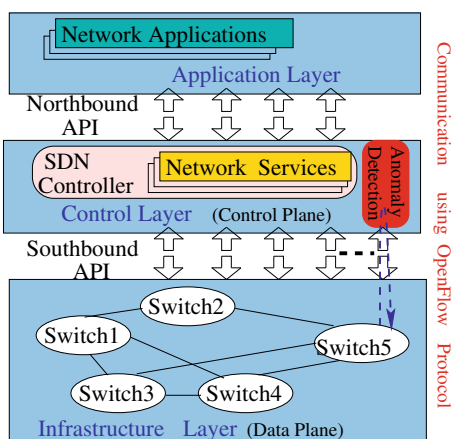


Figure 1 shows that the control plane act as a middleware between the network devices at the bottom and the network applications at the top. In SDN, flow management is a crucial task for making the network intelligent. The network intelligence by controlling the flows in the network is provided by the brain of the network, i.e., controllers. The switches and routers in SDN are not provided with any intelligence and they therefore simply accept rules from the controller for forwarding the packets. SDN mainly uses two interfaces, i.e., southbound and northbound interfaces. The southbound interface is used to pass information from the SDN controller to routers and switches. Similarly, the northbound interface is used to relay information from the SDN controller to applications and services that are running over the network. OpenFlow is a protocol used in SDN for providing centralized control functionality to the switches and traffic flows in a given network.

The growing popularity and recent advancements in SDN enable the inclusion of management plane along with control and data planes. The software services that can help in remotely monitoring and configuring the controller function are part of the management plane. It includes the definition of the policies of the network and its execution on the control plane, where data is forwarded by the data plane.

With the invention of the SDN, there is an architectural and structural modification in the traditional network. The logical centralization of SDN helps in dynamically adjusting the data traffic from a single point of control, i.e., controller. In SDN, the single point control is beneficial in providing the vertical decentralization of network and increasing the network efficiency but it is a vulnerable point for attack. The person who can get the access of the controller can demolish the entire network performance.

The intrusion detection in SDN is one of the critical attacks where less attention from the researcher is paid despite a huge amount of work done in the area of Openflow. The work emphasized on anomaly detection in SDN is covered in [4, 6, 7, 12]. In [4], the authors have used the Openflow architecture for detecting DDoS attack. In [12], the authors used various anomaly detection algorithms in experiments, where the algorithm was validated in both Small Office/Home Office (SOHO) and purely home networks. In [5, 6], the authors have used the OpenFlow protocol for enhancing the Remote Triggered Black Hole (RTBH) routing approach such that it can help to overcome the DDoS attack. The authors in [16] proposed an algorithm that is capable of dynamically changing the measurement granularity in both spatial and temporal dimensions for balancing the trade-off between error monitoring overhead and accuracy of anomaly detection. The author in [15] has elaborated attack scenarios and implementing them as SDN applications. They have used machine learning algorithms that are evaluated for their aptitude to detect anomalies in the SDN control plane.

- **OUR CONTRIBUTIONS:** In this paper, we consider different bandwidth of the network for detection of an anomaly in SDN. We also evaluate the performance of the SDN on the same network with different tree topologies. We have implemented tree topology on the OpenFlow controller using Mininet network emulator. We

considered OpenFlow messages as performance metrics for evaluating the performance of the network with different tree topologies.

The rest of the paper is organized as follows: In the next section, we discuss anomaly detection approaches in SDN. The experiment parameters, performance metrics, and results are given in Sects. 3 and 4 conclude this paper.

## 2 Anomaly Detection

The aim of the anomaly detection technique is to identify potentially harmful traffic in the computer network. Anomalies in the computer network are defined as *the patterns in data that do not resemble a well-defined notion of normal traffic flow in the network* [8, 9, 13, 15]. If an anomaly occurs because of malicious packets (e.g., originating from malware), inspecting the packet's payload is an effective way to recognize abnormal traffic. Two types of payload-based classifiers exist—Deep Packet Inspection (DPI) and Stochastic Packet Inspection (SPI). These methods provide very accurate results; however, the computational costs are high. Thus, approaches that merely need header fields instead of packet payload are required. However, building a strict model which is able to isolate the *normal* network traffic is very difficult. Hence, detecting anomalies in network traffic is a complex task. Traffic classification can also be employed to detect anomalies in a network.

In this paper, we are using rule-based traffic filtering as a subset of common protocols. We consider that the packet which has to pass through the centralized SDN controller must satisfy at least one filtering rule. In the case of filtering rule satisfaction, the flows are installed on the data plane. These installed rules help in packet forwarding inside the network in the future course of action without any involvement of the controller. In our experiment, we include non-IP packets and TCP packets. The controller computes an anomaly count based upon the arrival pattern of each byte of the data packet in the network. A predefined threshold is calculated for matching each packet against it. If the frequency of arrival of the packet is more than the threshold, then it is declared as an anomaly. Figure 1 illustrates a packet passing through the SDN controller. The anomaly detection process works with the SDN controller. The packet flows in the switch if the anomaly detection conditions are satisfied.

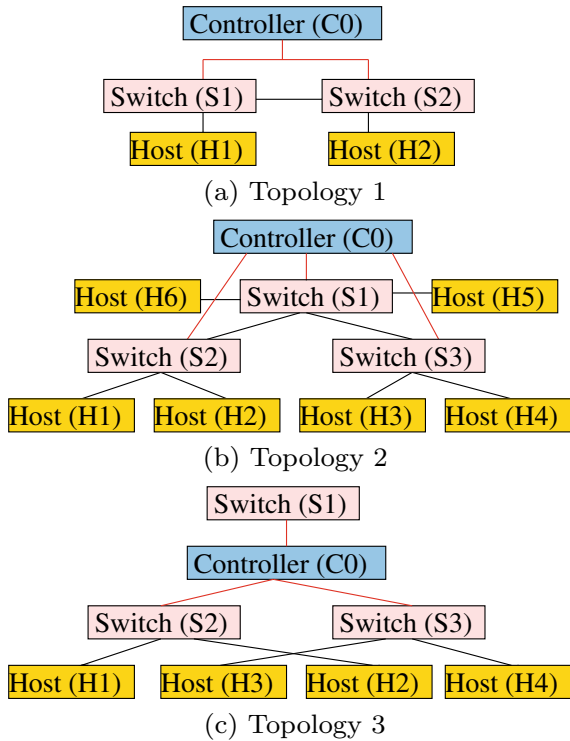
## 3 Experimental Results

In SDN, various types of OpenFlow messages have different impacts on user traffic. The accuracy and granularity of network flow measurement play an important role in anomaly detection in a network. In this paper, we investigate an easier and effective mechanism for implementing SDN in Mininet.

### 3.1 Experimental Setup on Mininet

In the proposed anomaly detection method, we used Mininet network emulator version 2.2.1 for creating network topologies in SDN [1]. Mininet is written in Python programming language and creates Open vSwitch version 2.0.2. We used POX controller that is a Python programming language based open source SDN controller. We created three tree topologies on Mininet. The first topology (Topology 1) consists of two OpenFlow virtual switches, two hosts, and a controller. The second topology (Topology 2) consists of three OpenFlow virtual switches, six hosts, and a controller. The third topology (Topology 3) consists of three OpenFlow virtual switches, four hosts, and a controller. Figure 2 illustrates all three topologies. The controller was made to run on a separate Internet Protocol (IP) address to avoid unnecessary traffic and for the proper capture of OpenFlow messages. We used *Wireshark* for capturing the traffic.

**Fig. 2** Illustration of different tree topologies that are considered for experimentation



### 3.2 Dataset and Traffic Generation

We are using CAIDA dataset [14] in order to generate network traffic that aims to evaluate our anomaly detection method. Since the dataset size was huge, we split the dataset into smaller files using *Wireshark*. We replaced the IP addresses in the dataset with the IP addresses of the hosts in Mininet topology using *bittwiste*. The dataset was replayed using *tcpreplay* [3] at different speeds in mbps. In order to delete the flow tables of the switches after each and every packet, we modified the POX controller [2] code so that the controller may receive all the packets in *l2\_learning.py*.

### 3.3 Impact on Packet Transmission

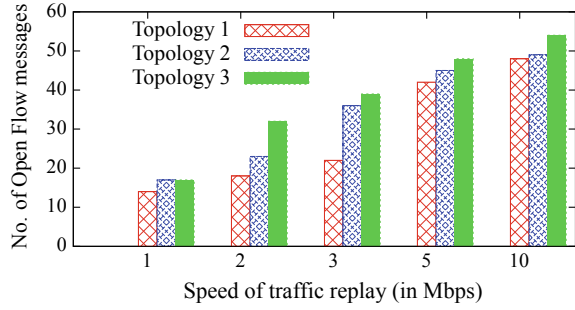
We run *tcpreplay* for different speeds in packet transmission and count the number of packets received by the controller. For anomaly detection, we have to use the minimum possible speed of traffic in order to check all the packets. We can get control messages for all the packets as the controller is deleting the flows after each packet. We set the bandwidth of topology links in *miniedit* at different bandwidths and calculated the number of packets received by the controller. We have considered three topologies as shown in Fig. 2. The packet dataset was replayed at different speeds using *tcpreplay*. The number of packets received was determined by using *Wireshark* and we observed that:

1. Below the set threshold bandwidth, the number of packets received increased by increasing the speed of replay.
2. Above the set bandwidths, the number of received packets remained almost the same.

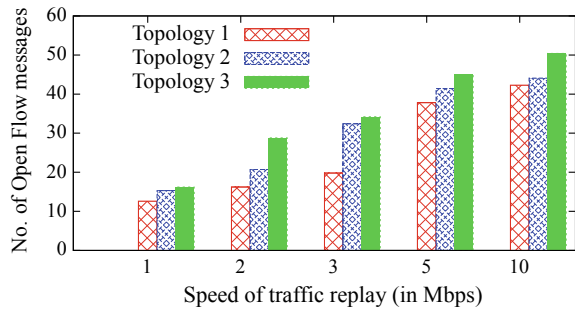
We used three different topologies for analysis which were with varying depths and complexities. The results illustrate that with the increase in the complexity of the topology, the number of received packets also increases for the same speed of replay. We used *iperf* command to cross-check whether the bandwidths were set with accuracy.

Figure 3 illustrates the impact of the network topologies (Topology 1, Topology 2, and Topology 3) and network speed on the OpenFlow messages. Part (a) and Part (b) of Fig. 3 illustrate that the OpenFlow messages count increases with an increase in the network data flow speed. The figure shows that Topology 1 requires less number of messages than Topology 3. This is because Topology 1 consists of less number of switches. The figure shows that the low network speed requires less number of OpenFlow messages. An interesting observation from the results is that less number of switches (i.e., Topology 1) require less OpenFlow messages and therefore the time complexity of anomaly detection will decrease.

**Fig. 3** Impact of network topologies and speed on the OpenFlow messages



(a) Network Speed 6Mbps.



(b) Network Speed 3Mbps.

## 4 Conclusion and Future Work

In this paper, we considered different network topologies to evaluate the performance of SDN. We have implemented the topologies on OpenFlow controller using the Mininet network emulator. We implemented the anomaly detection process on OpenFlow SDN controller. The process is written in Python programming language. We have used a standard database to evaluate the performance of SDN. We considered OpenFlow messages as performance metrics for evaluating the performance of the network. The results show that the number of OpenFlow messages increases with the speed of the network. We also observed that the network topology plays an important role in improving the performance of SDN. As a future work, we plan to detect anomalous packets by implementing machine learning algorithms in the SDN controllers. We will train our model on network dataset and merge this model with an SDN controller.

**Acknowledgements** This work is supported by the Science and Engineering Research Board (SERB) file number ECR/2016/000406/ES, project entitled as Development of an Energy-efficient Wireless Sensor Network for Precision Agriculture, and scheme Early Career Research Award.



## References

1. (2016) Mininet: An Instant Virtual Network on your Laptop. <http://mininet.org>.
2. (2016) POX: An Openflow controller. <http://www.noxrepo.org/pox/about-pox>.
3. AppNeta. (2016). Tcpreplay. <http://tcpreplay.synfin.net>.
4. Braga, R., Mota, E., & Passito, A. (2010). Lightweight DDoS flooding attack detection using NOX/OpenFlow. In *Proceedings of IEEE Conference on Local Computer Networks (LCN)*, pp. 408–415.
5. Giotis, K., Androulidakis, G., & Maglaris, V. (2014). Leveraging SDN for efficient anomaly detection and mitigation on legacy networks. In *Proceedings of European Workshop on Software Defined Networks*, pp. 85–90.
6. Giotis, K., Argyropoulos, C., Androulidakis, G., Kalogeras, D., & Maglaris, V. (2014). Combining OpenFlow and sFlow for an effective and scalable anomaly detection and mitigation mechanism on SDN environments. *Computer Networks*, 62, 122–136.
7. Gupta, H. P., Rao, S. V., & Tamarapalli, V. (2015). Analysis of stochastic  $k$ -coverage and connectivity in sensor networks with boundary deployment. *IEEE Transactions on Intelligent Transportation Systems*, 16(4), 1861–1871.
8. Gupta, H. P., Rao, S. V., & Venkatesh, T. (2016). Sleep scheduling protocol for  $k$ -coverage of three-dimensional heterogeneous wsns. *IEEE Transactions on Vehicular Technology*, 65(10), 8423–8431.
9. Gupta, H. P., Venkatesh, T., Rao, S. V., Dutta, T., & Iyer, R. R. (2017). Analysis of coverage under border effects in three-dimensional mobile sensor networks. *IEEE Transactions on Mobile Computing*, 16(9), 2436–2449.
10. Lange, S., et al. (2015). Heuristic approaches to the controller placement problem in large scale sdn networks. *IEEE Transactions on Network and Service Management*, 12(1), 4–17.
11. Lopes, F. A., Santos, M., Fidalgo, R., & Fernandes, S. (2016). A software engineering perspective on sdn programmability. *IEEE Communications Surveys Tutorials*, 18(2), 1255–1272.
12. Mehdi, S. A., Khalid, J., & Khayam, S. A. (2011). Revisiting traffic anomaly detection using software defined networking. I: *Proceedings of International Symposium on Recent Advances in Intrusion Detection (RAID)*, Springer Berlin Heidelberg, pp. 161–180.
13. Prabhu, G., & Jagatheesan, S. (2018) An efficient predictive network anomaly detection and visualization. *International Journal of Engineering Science*, 16651.
14. Singh, K. J., Thongam, K., & De, T. (2018). Detection and differentiation of application layer ddos attack from flash events using fuzzy-ga computation. *IET Information Security*, 12(6), 502–512.
15. Sommer, V. (2014). Anomaly detection in the SDN control plane. Master’s thesis, Technische Universität München.
16. Zhang, Y. (2013). An adaptive flow counting method for anomaly detection in SDN. In *Proceedings of ACM Conference on Emerging Networking Experiments and Technologies*, pp. 25–30.

# An Extended Playfair Encryption Technique Based on Fibonacci Series



Mohd Vasim Ahamad, Mohd Imran, Nazish Siddiqui and Tasleem Jamal

**Abstract** The rapid advancement in networking technologies leads to the sharing of information by millions of users at a much higher rate. Information sharing over unprotected network may lead to compromise of the sensitive and confidential information to unauthorized person. Cryptography is one of the several ways to guard sensitive and confidential information. Before sharing the information over the Internet, it must be encrypted to preserve its confidentiality. Encryption is a process of converting readable information into scrambled form so that no unauthorized person can make use of it. In this paper, the Playfair encryption technique is taken into consideration for encrypting the information to be shared. In its simplest form, the Playfair encryption technique generates a  $5 \times 5$  key table by taking a key as input and encrypts the diagraphs of actual message. The key generation process is improved by modifying it with Fibonacci series. The Fibonacci series is used create a random key, which is passed to generate  $5 \times 5$  key table. This key table is used to encrypt the actual message using rules defined by Playfair encryption algorithm. In this paper, the limitations of  $5 \times 5$  key table are removed by using an  $8 \times 8$  key table, which provides much higher level of security to the message being encrypted.

**Keywords** Encryption · Playfair encryption · Cryptography · Information security · Fibonacci series for encryption

---

M. V. Ahamad (✉) · N. Siddiqui  
Department of Computer Science & Engineering, University Polytechnic, Integral University,  
Lucknow, India  
e-mail: [vasim.iu@gmail.com](mailto:vasim.iu@gmail.com)

N. Siddiqui  
e-mail: [nazishcs016@gmail.com](mailto:nazishcs016@gmail.com)

M. Imran  
Department of Computer Engineering, ZHCET, Aligarh Muslim University, Aligarh, India  
e-mail: [mimran.ce@amu.ac.in](mailto:mimran.ce@amu.ac.in)

T. Jamal  
Department of Information Technology, REC, Azamgarh, India  
e-mail: [tasleemjamal51@gmail.com](mailto:tasleemjamal51@gmail.com)

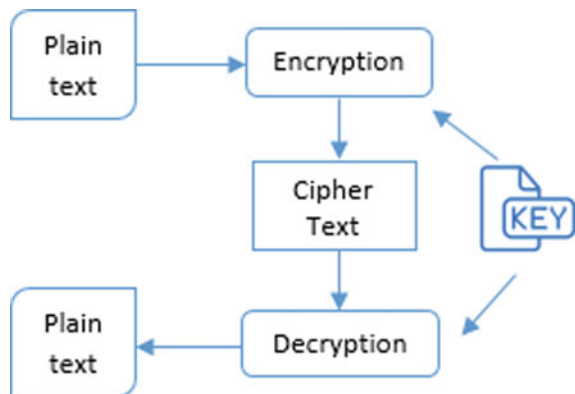
## 1 Introduction

The recent evolution in communication and internet technologies results in sharing of information among users in a secure way in the form of files, emails, e-commerce activities, transaction, etc. [1, 2]. To share the confidential information over such an unprotected medium, users need to be ensured that no unauthorized user can read this information. To ensure the security of private or confidential information, encryption techniques can be used. Encryption can be defined as the method of converting the readable message into unreadable or scrambled message [3]. In cryptography, the encryption method inputs the original message to an encryption algorithm and converts the message into an unreadable form using a secret key [4, 5]. This unreadable or scrambled message is called as the ciphertext [6]. The authorized recipient can recover the original message by decrypting the ciphertext with the help of the secret key [7, 8]. A secret key, or simply a key, is a combination of alphanumeric text optionally having special symbols too [9, 10]. The secret key is given as input to the encryption algorithm with plain text as another input. With the help of the secret key, the plain text is transformed into the ciphertext. It is also decrypting the ciphertext to recover the original message. The study and application of encryption and decryption are called as the cryptography. Figure 1 shows how a message is encrypted to produce ciphertext and decrypted to recover the original message.

## 2 Related Work

In [11], the authors analyzed and modified the Playfair encryption algorithm. They used 8\*8 key table to overcome the limitations of  $5 \times 5$  key table and used LFSR to generate the random numbers. The authors in [12] modified Playfair encryption with Fibonacci series. They converted the plain text into the ciphertext by interchanging each character in plain text with Fibonacci number generated characters. In [4], the

**Fig. 1** The encryption process



authors have modified the Playfair encryption with Fibonacci series. They generated a random key of fixed length by generating the next six terms with respect to the input Fibonacci term. Further, the key generated by the previous step is placed into  $5 \times 5$  key table and original message is encrypted using the Playfair encryption rules. In [13], the authors used  $6 \times 6$  matrix covering 26 alphabets and digits from 0 to 9, instead of using  $5 \times 5$  which covers only 26 alphabets. There are many versions of Playfair encryption algorithms are being used for protecting confidential information. The motivation behind this work is to provide more security to confidential information. To achieve this, the Playfair Encryption method is modified with the help of Fibonacci Series generated secret key. The Playfair encryption algorithm needs a secret key to generate a  $5 \times 5$  key table, which transforms the plaintext into ciphertext. This secret key is generated with the help of Fibonacci series. To generate more complex secret key, the  $8 \times 8$  key generation table is used which accommodate 26 uppercase, 26 lowercase alphabets, 10 digits, and two special symbols.

### 3 Playfair Encryption Technique

Encryption is a method for transforming the original and readable message into a scrambled message. Substitution and transposition are the two basic approaches to encryption techniques. In substitution based encryption, letters and other symbols of plaintext are replaced by other letters, digits or special symbols. The substitution based encryption techniques involve the replacement of a plain text symbol with a ciphertext symbol. Examples of substitution based encryption techniques are Caesar cipher, Playfair, Hill cipher, and One-Time Pad. In transposition-based encryption, the mapping of the plaintext with the cipher text is done by performing some sort of permutation on the plaintext letters.

Playfair encryption technique is one of the approaches used in substitution-based encryption technique [6]. Unlike simple substitution methods, which encrypt single letter at a time, Playfair encrypts pair of letters at a time. Playfair encryption is a kind of block cipher which can encrypt message with no limit on the number of characters [6, 7]. It encrypts and decrypts digrams (two characters) at a time. The cryptanalysis on Playfair needs to break 600 possible digraphs in comparison with simple substitution, which requires to break just 26 possible options [8]. The first step of Playfair encryption is to generate a  $5 \times 5$  key table having 26 uppercase letters (I/J is placed in a single column). To generate the key table, a secret key is selected. This secret key helps to encrypt the digrams using Playfair encryption. While generating the key table, each non-repeating letter of secret key is placed in a separate cell of the  $5 \times 5$  grid starting from the left top of the grid [8, 9]. Then remaining letters are placed in alphabetical order. Table 1 is demonstrating the key table generation with the secret key “CRYPTO”.

**Table 1** 5×5 Key table

C	R	Y	P	T
O	A	B	D	E
F	G	H	I/J	K
L	M	N	Q	S
U	V	W	X	Z

Now, the key table is generated and is ready to encrypt the plaintext in the form of digrams. Let us assume, the sender wants to communicate the message “SECRET WORK” to the receiver. Firstly, this message is converted into the blocks of two characters (digrams) as “SE CR ET WO RK”. It is clear from Fig. 1 that an encryption algorithm needs plaintext and a key to generate the ciphertext. Figure 2 is demonstrating the steps of Playfair encryption technique which accepts the key table and digrams of plaintext and converts it to ciphertext. Using the Playfair encryption rules, the digrams of the plaintext can be transformed into the ciphertext with the help of key table. The Fig. 3a–e showing the ciphertexts of each digram using the Playfair encryption algorithm. The plaintext digrams “SE CR ET WO RK” is encrypted using Playfair encryption rules as shown in Fig. 3.

In Fig. 3a, both the letters of digram “SE” belongs to the same column, and hence first rule is applied and the ciphertext “ZK” is generated. In Fig. 3b, letters “CR” falls in the same row. Here rule 2 is applied and “RY” is generated as ciphertext. Figure 3c, b, because the letters “ET” belong to the same row. Following the rule 2, “TZ” becomes the ciphertext. Figure 3d, e follow rule 3, because the letters in

1. If letters of the digram come under the same column, replace with the letter beneath. If a letter of the digram is placed at the bottom of the key table, it is replaced by the letter circularly following the topmost in the same column.
2. If both the letters of the digram falls into the same row, replace letters of digram with the just right of each letter. If one of the letters in digraph is placed at rightmost cell of the key table, then replace it with the letter in leftmost cell in the same row.
3. Otherwise, a rectangle is created with the two letters of the digram, and each letter is replaced with the letters on the horizontal opposite corner of the same row.

**Fig. 2** Rules of Playfair encryption

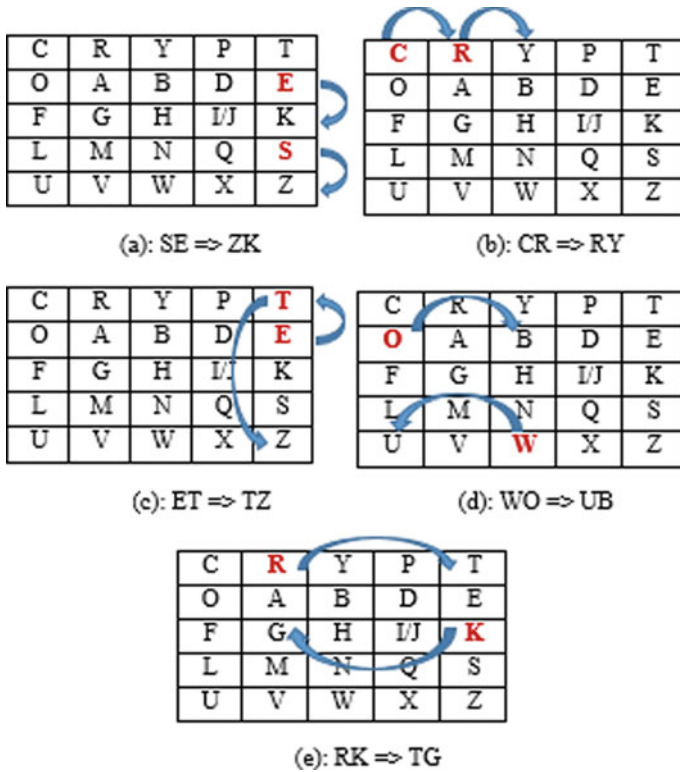


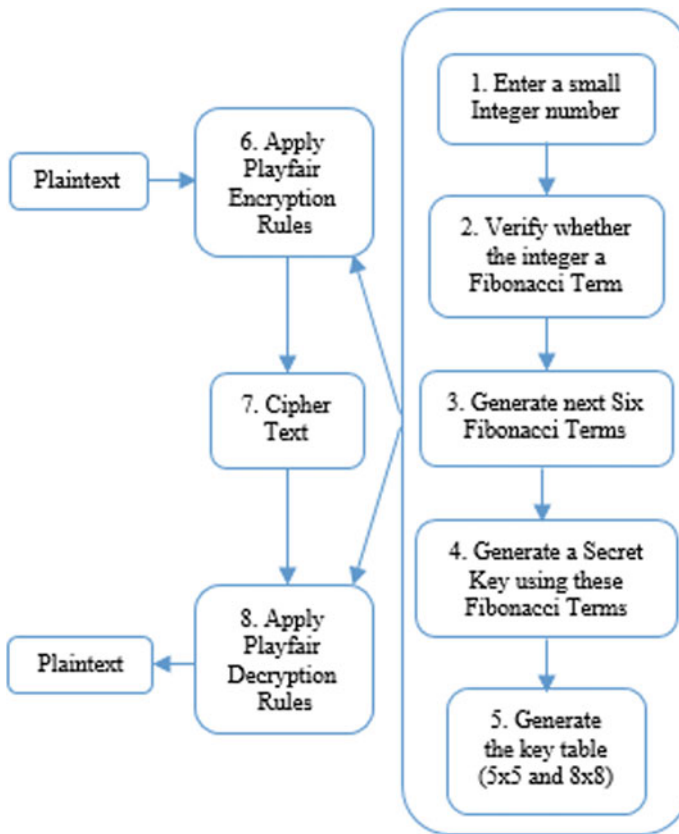
Fig. 3 a–e Playfair encryption example

digrams “WO” and “RK” form a rectangle. Using rule 3, the ciphertexts “UB” and “TG” are generated, respectively. So, the plaintext “SECRET WORK” is encrypted using the Playfair encryption technique and the ciphertext “ZKRYTZ UBTG” is generated.

### 4 Proposed Work

In this paper, the goal is to use the Playfair encryption algorithm to encrypt the textual messages. To encrypt a message (plaintext), it is provided as input to the Playfair encryption algorithm with a secret key as another input. The Playfair encryption is a manual symmetric key encryption algorithm which allows sharing the same secret key to both the sender of the message and the receiver.

Generally, the secret key in the Playfair encryption algorithm is selected beforehand. It can cause the selection of very common secret key or compromise of secret key while sharing to the receiver end. To eliminate these issues, this paper aims to



**Fig. 4** The proposed algorithm

generate a secret key using Fibonacci series terms. It will generate the random secret keys each time the algorithm is executed. The proposed method comprises three stages namely the key generation, the encryption, and the decryption. The proposed method can easily be visualized in Fig. 4, in which rightmost box (step 1–5) shows the key generation step. The encryption algorithm (step 6) takes the generated secret key and plaintext as input and generates ciphertext (step 7). The ciphertext and the secret key is provided to the decryption algorithm (step 8), which recovers the plaintext.

#### **4.1 The Key Generation Using $5 \times 5$ Key Table**

The Playfair encryption, like other encryption algorithms, needs a secret key to encrypt the plaintext provided as input. Generally, in simple Playfair encryption, the secret key is selected beforehand manually. In this paper, the secret key is generated

on the run using Fibonacci series terms. The steps to generate the secret key is discussed as follows.

- Select a small integer number.
- Check this number for the valid Fibonacci term, if not, provide error message.
- If it is a Fibonacci term, generate the next six terms of the Fibonacci Series.
- Once the Fibonacci terms are generated, convert them into ASCII range of alphabets.
- These seven alphabets become the secret key for the Playfair encryption.

Once the secret key is generated, it is placed into the  $5 \times 5$  key table as discussed in Sect. 3 and shown in Table 1.

### ***4.2 The Encryption Using $5 \times 5$ Key Table***

The Playfair encryption with the help of  $5 \times 5$  key table is discussed in Fig. 2 of Sect. 3.

### ***4.3 The Decryption Using $5 \times 5$ Key Table***

The Playfair cryptography approach is a manual symmetric approach which considers the same secret key for encryption and decryption. In this approach, the decryption approach is just the opposite of the encryption algorithm. The steps of the decryption algorithm are discussed in Fig. 5.

### ***4.4 The Key Generation Using $8 \times 8$ Key Table***

The main limitation of constructing  $5 \times 5$  key table is that it considers I and J as one character, which affects the encryption and decryption. It accommodates only capital letters in  $5 \times 5$  key table. Due to this limitation, the only simple combination of capital letters can be used while generating the secret key. The  $5 \times 5$  key table also lacks to include digits which strengthen the encryption of plaintext. Another limitation of  $5 \times 5$  key table is that it doesn't incorporate space. To deal with the above-mentioned issues, this works also includes the generation of  $8 \times 8$  key table. It contains 26 lowercase letters, 26 uppercase letters, 10 numerical digits, underscore (\_) to represent space, and \$ to accommodate the repeating alphabets. Considering "Crypto123" as the secret key, the  $8 \times 8$  key table can be generated as follows.



1. If letters of the digram come under the same column, replace with the letter above each one. If a letter of the digram is placed at the top of the key table, it is replaced by the letter circularly following the bottom in the same column.
2. If both the letters of the digram falls into the same row, replace letters of digram with the just left to each letter. If one of the letters in digraph is placed at leftmost cell of the key table, then replace it with the letter in rightmost cell in the same row.
3. Otherwise, a rectangle is created with the two letters of the digram, and each letter is replaced with the letters on the horizontal opposite corner of the same row.

**Fig. 5** The decryption process

#### ***4.5 The Encryption Using $8 \times 8$ Key Table***

The primary steps of encryption algorithm using  $8 \times 8$  key table is similar to that of  $5 \times 5$  key table as discussed in Fig. 2. The following additional steps are used when space and duplicate letters are encountered in the plaintext.

- If a digram contains a letter and a space, put underscore ( ) in place of the space. (e.g. Plaintext: ENCRYPT IT, Digrams: EN CR YP T\_ IT)
- If a digram contains duplicate letters, put the dollar sign (\$) in place of the last letter. (e.g. Plaintext: GREETINGS, Digrams: GR E\$ TI NG S\_)

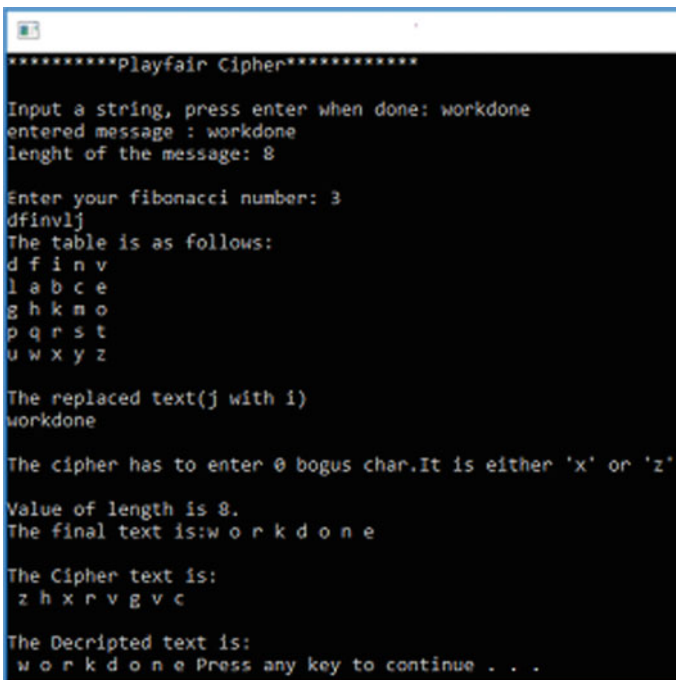
#### ***4.6 The Encryption Using $8 \times 8$ Key Table***

The initial steps of decryption algorithm using  $8 \times 8$  key table is similar to that of  $5 \times 5$  key table as discussed in Fig. 5. The following additional steps are used when space and duplicate letters are encountered in the ciphertext.

- If a digram of ciphertext contains underscore ( ), it is replaced by space.
- If a digram of ciphertext contains \$ (dollar sign), then replace it with just a previous letter recovered from cipher text digram.

## 5 Results and Discussion

The Playfair encryption algorithm receives two inputs namely the secret key and the plaintext to be encrypted. The extended Playfair algorithm discussed in this paper generates the secret key using Fibonacci terms as discussed in Sect. 5. The secret key is placed into the  $5 \times 5$  key table according to the rules discussed in Table 1 of Sect. 4. In addition to the secret key, the plaintext is supplied to the extended Playfair encryption algorithm and the ciphertext is generated. The results of all steps are shown in the Fig. 6. To deal with the limitations of  $5 \times 5$  key table, the  $8 \times 8$  key table is introduced which contains uppercase and lowercase letters, digits, underscore, and dollar sign. This helps in creating more complex secret key and strengthen the encryption approach. The secret key consisting of uppercase and lowercase letters, and digits is generated and placed into an  $8 \times 8$  key table according to the rules discussed in Table 2. Figure 7 shows the results of each step using  $8 \times 8$  key table.



```
*****Playfair Cipher*****
Input a string, press enter when done: workdone
entered message : workdone
lenght of the message: 8

Enter your fibonacci number: 3
dfinvlj
The table is as follows:
d f i n v
l a b c e
g h k m o
p q r s t
u w x y z

The replaced text(j with i)
workdone

The cipher has to enter 0 bogus char.It is either 'x' or 'z'
Value of length is 8.
The final text is:w o r k d o n e

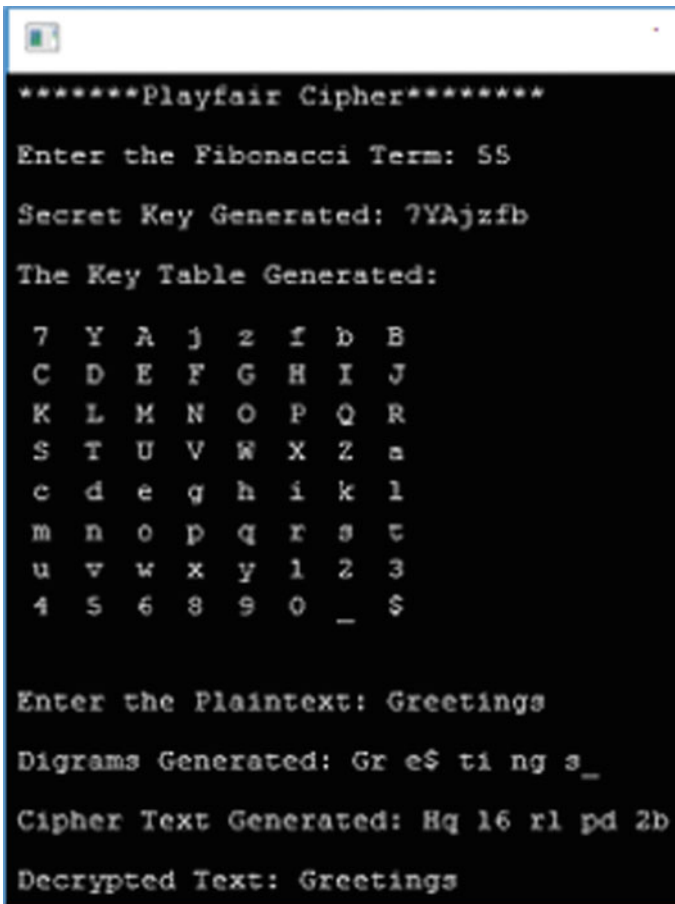
The Cipher text is:
z h x r v g v c

The Decrypted text is:
w o r k d o n e Press any key to continue . . .
```

Fig. 6 Playfair encryption and decryption using  $5 \times 5$  Key table

**Table 2** The  $8 \times 8$  Key table

C	r	p	t	o	l	2	3
A	B	D	E	F	G	H	I
J	K	L	M	N	O	P	Q
R	S	T	U	V	W	X	Y
Z	a	b	c	d	e	f	g
h	i	j	k	l	m	n	q
s	u	v	w	x	y	z	4
5	6	7	8	9	0	_	\$



**Fig. 7** Playfair encryption and decryption using  $8 \times 8$  Key table

## 6 Conclusion and Future Directions

Cryptography is a study and application of methods to secure the information shared through a potentially vulnerable medium. It also ensures that no unauthorized person can get the information being shared. Encryption is a method for transforming the plaintext into a scrambled message (ciphertext). Playfair encryption technique is a kind of block cipher which can encrypt message in the form digrams (two characters) at a time. The Playfair encryption algorithm needs the secret key and the plaintext as inputs. The approach discussed in this paper generates the secret key using Fibonacci terms so that random and secure secret key can be generated. The secret key is placed into  $5 \times 5$  key table and the plaintext is provided to the extended Playfair encryption algorithm to generate the ciphertext. The limitation, of not having the combination of uppercase & lowercase letters, digits and special symbols in  $5 \times 5$  key table, the need for  $8 \times 8$  key table arises. It helps in generating more complex secret key and strengthen the encryption approach. The secret key consisting of uppercase and lowercase letters, and digits is generated and placed into an  $8 \times 8$  key table. It also incorporates the underscore symbol to represent space, and \$ to accommodate the repeating alphabets. Using the extended Playfair encryption technique, one can generate stronger secret key and accommodate the spaces and repeating letters in plaintext. Researchers can use other efficient methods instead of the Fibonacci terms to generate stronger secret keys. Another improvement in this direction can be the inclusion of other special symbols, and punctuation marks.

## References

1. Akhtar, N., & Ahamad, M. V. (2017). Graph tools for social network analysis. In N. Meghanathan (eds.) *Graph Theoretic Approaches for Analyzing Large-Scale Social Networks*, pp. 18–33. IGI Global.
2. Khan, H., Ahamad, M. V., & Samad, A. (2017). Security challenges and threats in cloud computing systems. *International Journal of Advanced Research in Computer Science*, 8(2), 36–39.
3. Rivest, R. L. (1990). Cryptography. In Van Leeuwen, J. (ed.) *Handbook of Theoretical Computer Science*, vol. 1. Elsevier.
4. Ahamad, M. V., Masroor, M., & Fatima, U. (2017). A modified Playfair encryption using fibonacci numbers. *International Journal of Advanced Technology in Engineering and Science*, 5(6), 306–315.
5. Khan, M. A., Mishra, K. K., & Jayakumari, S. J. (2015). A new hybrid technique for data encryption. In *Proceedings of Global Conference on Communication Technologies*.
6. Playfair Cipher. (2017). [https://en.wikipedia.org/wiki/Playfair\\_cipher](https://en.wikipedia.org/wiki/Playfair_cipher).
7. Goyal, P., Sharma, G., & Kushwah, S. S. (2015). A new modified Playfair algorithm using CBC. In *International Conference on Computational Intelligence and Communication Networks*, pp. 1008–1012.
8. Srivastava, S. S., & Gupta, N. (2011). A novel approach to security using extended Playfair cipher. *International Journal of Computer Applications*, 20(6), 39–43.
9. Bhattacharyya, S., Chand, N., & Chakraborty, S. (2014). A modified encryption technique using Playfair Cipher 10 by 9 matrix with six iteration steps. *International Journal of Advanced Research in Computer Engineering & Technology*, 3(2), 308–312.

10. Dipthi, R., A survey paper on Playfair cipher and its variants. *International Research Journal of Engineering and Technology*, 4(4), 2607–2610.
11. Sinkov, A. (1996). Elementary cryptanalysis: A mathematical approach. *Mathematical Association of America*.
12. Agarwal, P., Agarwal, N., & Saxena, R. (2015). Data encryption through fibonacci sequence and unicode characters. *MIT International Journal of Computer Science and Information Technology*, 5(2), 79–82.
13. Kumar, A., Mehra, P. S., Gupta, G., & Sharma, M. (2013). Enhanced Block Playfair cipher. In *International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness*, pp. 689–695.

# An Automated System for Epileptic Seizure Detection Using EEG



Bilal Alam Khan, Anam Hashmi and Omar Farooq

**Abstract** Epileptic seizures are usually investigated using EEG. The dynamic and statistical properties of brain waves of an individual with seizure are different from a normal person's brain waves. This paper exploits these underlying properties of EEG using Lyapunov exponent and approximate entropy and proposes a novel statistical feature namely Gini's coefficient. In this paper, we propose an automated system for detecting seizure using statistical and machine learning algorithm. The data used was publicly available with five different classes (normal to seizure). Linear discriminant analysis (LDA) was used to classify the extracted features. The proposed method gives the best accuracy of 100% in detecting seizure from the EEG.

**Keywords** Epileptic seizure · EEG · LDA · Gini's coefficient · Machine learning

## 1 Introduction

Epilepsy is one of the most common and debilitating neurological disorders. It is described by periodic and unprovoked transient disturbances of perception resulting from the excessive synchronous discharge of neurons in the brain [1]. According to WHO, 0.6–0.8% of the world population is affected by epilepsy and almost 80% of the affected people are found in developing countries [2, 3]. The electroencephalogram (EEG) is the most widely used and recognized technique for the investigation of brain signals in general and epileptic seizures in particular [4]. EEG is usually performed in the neurophysiological laboratory for short period analysis of brain waves. However, for better quality, completeness, and comprehensiveness of the data, the observation

---

B. A. Khan (✉) · A. Hashmi · O. Farooq  
Department of Electronics Engineering, Aligarh Muslim University, Aligarh 202001,  
UP, India

e-mail: [bilalalamwaris@gmail.com](mailto:bilalalamwaris@gmail.com)

A. Hashmi

e-mail: [anamhashmi360@gmail.com](mailto:anamhashmi360@gmail.com)

O. Farooq

e-mail: [omar.farooq@amu.ac.in](mailto:omar.farooq@amu.ac.in)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_15](https://doi.org/10.1007/978-981-15-0694-9_15)

of seizure activity of the human brain by multi-channel EEG is done over a long period (more than 24 h) in critically ill patients [5]. Thus, the size of the generated data is enormous and this makes the analysis labor-intensive, time-consuming, cumbersome, and error prone. Even the most qualified neurologists find it challenging to identify seizure because of the enormity of the data and because of the added ocular and muscular artifacts [6].

To mitigate this problem and to help the well-trained healthcare providers, seizure detection using computer algorithms has become an area of interest; especially the field of analyzing signals using machine learning tools has caught the attention of various new researchers.

In the past few decades, a lot of work has been done in the field of seizure detection, different algorithms have been developed, new feature extraction strategies have been used and new methodologies are proposed as well. These include time, frequency, and time–frequency domain analysis. Amplitude, sharpness, and duration fall under the category of Time-based features. To exploit the frequency-based characteristics of the signal, time domain signals were mapped into the frequency domain using techniques that include fast Fourier transform, power spectral density, etc. [7, 8]. However, there is a shortcoming to these methods; the assumption that the EEG signals are stationary. EEG signals are naturally nonstationary, thus challenging the assumption [8]. Hence, a new method was developed where the signal is divided into windows of equal length and exploits both time domain as well as frequency domain features of EEG [9]. This method was termed as short-time Fourier transform (STFT). The window used in this method should be sufficiently small enough to make the assumption of stationarity applicable. Alongside these methods, it has been shown that EEG has nonlinear characteristics and this particular characteristic has received significant consideration as well by the community of neuro-scientists [10]. Predominantly, empirical mode decomposition (EMD), Fractal dimensions and entropy have been assumed to study and extract those underlying nonlinear attributes of the EEG which can be used for seizure detection [8].

In this study, we propose a hybrid model in which the advantages of the above-mentioned models are adopted while eliminating their shortcomings for the detection of an epileptic seizure. In addition, we propose the use of a novel feature, Gini's coefficient, for extracting distinguishing information from the EEG.

The structure of this paper is as follows. A short account of the data used in this work was provided in Sect. 2, where feature extraction techniques used were expounded after which the explanation of chosen features and clarification of the classification stage used in this study. Section 3 presents a brief summary of the experimental results, followed by the comparison with standard published work. Lastly, Sect. 4 describes the conclusion of this work.

## 2 Materials and Methodology

### 2.1 Data Used

The data used in this study was taken from publicly available and provided by Andrzejak et al. [11], University of Bonn [11]. The dataset comprises of 500 segments of 23.6 s each and is equally distributed into groups A, B, C, D, and E. These groups comprised of normal, inter-ictal, and ictal EEG signals recorded using a single channel. The placement of the electrodes was done in accordance with the standard 10–20 electrode placement system. EEG recorded over healthy volunteers was contained in set A and set B. The EEG recording of these volunteers was performed with opened eyes in set A and with closed eyes in set B while they were relaxing. Sets C, D, and E consisted of EEG recordings of five patients with a background of having seizure activity. Set C was recorded over the hippocampal zone and set D consisted of EEG recorded over the epileptogenic zone of the brain. However, both sets C and D have EEG recording from non-seizure (inter-ictal) interval. EEG segment recorded during seizure (ictal) activity is contained in set E. For creating a database, EEG was recorded using a 128-channel system and stored on a disk with a 12-bit analog to digital converter. The original analog EEG signal was sampled at a frequency of 173.61 Hz and therefore, the corresponding bandwidth according to Nyquist criteria would be 86.85 Hz.

### 2.2 Features

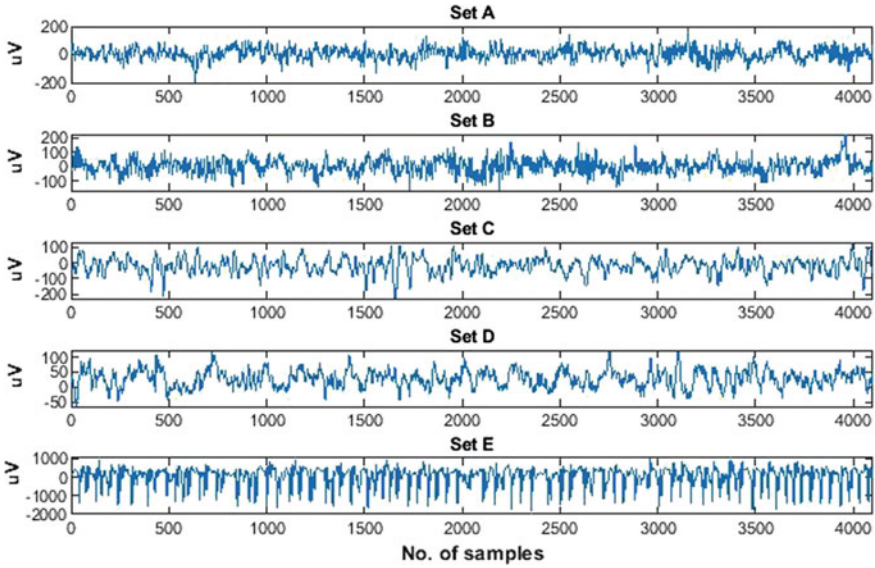
Figure 1 shows EEG signals of five different classes (A–E) based on normal, inter-ictal, and ictal criteria. Set A and set B correspond to the normal class. EEG waveforms of these two appear to be random and do not have any specific patterns while EEG waveforms of set E is recurring and follows a specific pattern. These distinctive properties of the data used, make the choice of approximate entropy natural to this study. Lyapunov exponent has also been considered in this study because of the dynamic characteristics of EEG signals before and after the seizure are known to be chaotic. Furthermore, the use of a novel statistical feature, namely, Gini's coefficient has been proposed.

In addition, it is known that EEG is a nonstationary signal and thus its statistical properties changes with time. However, windowing the EEG into small segments makes the assumption stationarity applicable. Therefore, the technique of windowing the signal into small windows is employed. Additional explanation of these features and the methodology used has been mentioned in the following sections.

#### **Lyapunov Exponent**

Lyapunov exponent is a numerical measure for differentiating amidst the several kinds of trajectories relying on the initial conditions [12]. It is a measure of the degree at which the orbits diverge one from other. Chaotic systems exhibit aperiodic





**Fig. 1** EEG waveforms showing different classes

behavior because phase space trajectories diverge at an increasingly high rate, which is equivalent to the exponential function. A negative exponent denotes the advancement of the trajectories to a mutual stationary point; a zero exponent implies that the trajectories are on a stable attractor. Finally, chaotic attractor trajectory points are suggested by a positive exponent [13]. There are two methods of calculating Lyapunov exponent, either from the equation of motion of the dynamic system or from the time series itself using local Jacobi matrices. The first one provides the approximation for the largest Lyapunov exponent (LLE) only. The second method can provide all the Lyapunov exponents. This study uses the algorithm by Wolf et al. to find the LLE [14].

The algorithm is described as follows.

Considering recorded EEG ( $y(t)$ ) as the time series in discussion and are mapped to the phase space. It is a specified time series with  $k$ -dimensions and time axis  $t$  and is given by

$$y(t), y(t + 1.t), \dots, y(t + (k - 1)t) \quad (1)$$

The location of the nearest neighbor with respect to the initial point:

$$y(t_0), y(t_0 + 1.t), \dots, y(t_0 + (k - 1)t) \quad (2)$$

$L(t_0)$  gives the initial separation between two points. After some time  $t_1$ , the length changes to  $L(t_1)$ . The average logarithmic rate of divergence of two initially neighboring trajectories is given by

$$\lambda = \frac{1}{t_m - t_o} \sum_{k=1}^M \log_2 \frac{L(t_k)}{L(t_k - 1)} \quad (3)$$

EEG signals corresponding to the normal (inter-ictal) recording is known to be random, dynamic, and chaotic. However, EEG recording of the ictal state is regular, less random, and thus lose chaotic nature. Thus, the choice of Lyapunov exponent appears to be natural in the analysis of EEG signals pertaining to the seizure activity.

### Approximate Entropy (ApEn)

Entropy is the measure of the rate of information or randomness [12]. Usually, high variability or randomness corresponds to a high value of entropy while an increased symmetry gives a low value of entropy. It is broadly categorized into embedding and spectral entropy. Spectral entropy is calculated from the spectrum of the signal while Embedding entropy are calculated directly from a time series, i.e., Kolmogorov–Sinai entropy and approximate entropy etc. [13]. In this study, approximate entropy has been used to extract relevant features.

Approximate entropy is a technique that is used to describe the unpredictability of both deterministic as well as stochastic signals [15]. Approximate entropy will have a higher value for irregular time series than one with symmetrical patterns. It is estimated by relating the similarity of the samples by pattern length ( $m$ ) and similarity coefficient ( $r$ ). The formula of Approximate entropy is shown mathematically:

$$ApEn = \ln (Cm(r)/Cm(r + 1)) \quad (4)$$

where

$C_m(r)$  is the pattern mean of length of  $m$

$C_m(r + 1)$  is pattern mean of length  $m + 1$ .

$$C_m(r) = \frac{n_m(r)}{N - m + 1} \quad (5)$$

### Gini's Coefficient

The Gini Index, also known as the Gini Coefficient, is used to measure inequality in wealth distribution and is still studied in relation to wealth distribution as well as in other areas.

As described by Hurley and Rickard, “Inequality in wealth” in signal processing language is “efficiency of representation” or “sparsity” [16].

In this study, Gini's coefficient is calculated using a method that is equivalent to the Lorentz curve definition. The method used is the relative mean absolute difference, which is a measure of statistical dispersion. It is equal to the average absolute difference of all pairs of samples in a data, taking the sum of these differences and normalizing them with the twice of the average [17].

Mathematically, Gini's coefficient ( $G$ ) is given by

$$G = \frac{\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|}{2 \sum_{i=1}^n \sum_{j=1}^n x_j} \quad (6)$$

where,

$x^i$  denotes the  $i$ th sample in data.

It is a very common feature in economics to study wealth distribution; however, its use in the signal-processing field has been very limited. Gini's coefficient is primarily used to study the sparsity of the signal. However, in this study, it has been used only as a time domain feature.

### 3 Methodology

The EEG data used consisted of five groups of 100 segments of 23.6 s. The sampling frequency was 173.61 Hz, thus producing signal 4097 ( $23.6 \times 173.61$ ) data points. These segments are distributed over five groups namely A, B, C, D and E. Each group consisted of 100 files, each of 4097 data points. Thus, the size of the matrix corresponding to each group will be  $[100 \times 4097]$ .

In this study, a window of 1 s was taken and was slid over the entire length of the signal. For each second window, 173 data points are analyzed and three features were calculated on them and the entire process is performed until the end of the signal length. The three features that were extracted include Lyapunov Exponent, Approximate Entropy, and Gini's coefficient. This process was performed for all the data in a group and thus a matrix of size  $[23 \times 100]$  for each feature, is obtained.

For each column of that matrix, the median was calculated which provided a better summary of the data. This gave us our feature vector of size  $[23 \times 1]$ . Thus combining the feature vectors a matrix is obtained of size  $[23 \times 3]$ . This whole process is repeated over five sets (A–E). These feature vectors are then divided into training and test set. The distribution of training set and test set comprised of 70% and 30% of features respectively. Fourfold cross-validation was used to reduce overfitting. Training data was used to train the classifier. After that, the trained classifier classified test set into their respective classes.

The nonlinear feature Gini's coefficient has not been used for the seizure detection to the best of our knowledge. In addition to this, the median also provides robustness to this method.

### 4 Classification

In this study, Linear Discriminant Analysis (LDA) based classification was used. The notion of LDA is to map feature vectors 'J' from a  $p$ -dimensional space to vector 'K' in a  $q$ -dimensional space using linear transformation so that the separation between the classes becomes maximum. Scatter matrices are used to formulate optimization. The most widely used optimization for LDA is

$$J_1(m) = tr\left(S_{2y}^{-1} S_{1y}\right) \tag{7}$$

$$J_2(m) = det\left(S_{2y}^{-1} S_{1y}\right) \tag{8}$$

where  $tr(A)$  denotes the trace of the matrix and  $det(A)$  denotes the determinant of the matrix, and  $S_{iy}$  is the scatter matrix in dimension  $y$ -space [18].

LDA was first trained on the training data and it was used to fit the training data according to the above-described method. Then the test data, which was the part of the same distribution over which LDA was trained, was tested for classification using the same.

## 5 Results

After the classification was performed using LDA, accuracy was calculated as a metric to quantify the capturing potential of the various features used. The classification accuracy has been reported in Table 1. The division of Table 1 is in accordance with the division prepared by Alam and Bhuiyan and followed by Khan and Farooq [19, 20]. The table has been divided into five cases, at the start; both case 1 and case 2 describe the binary classification. However, with a modification that case 1 describes classification between normal, set A and seizure affected (ictal) person, set E while case 2 describes classification between inter-ictal, set D and seizure (ictal), set E. Furthermore, case 3 and case 4 both corresponds to three-level classification. Case 3 shows classification between normal, inter-ictal and ictal. This case 3 corresponds to the classification between set A, set D and set E. Case 4 shows the similar classification but with a subtle difference that sets of healthy subjects (opened eyes and closed

**Table 1** Comparison of results between this work and the work done by Alam and Bhuiyan [19] as well as by Khan and Farooq [20]

Classifiers	Results from Alam and Bhuiyan [19]		Results from Khan and Farooq [20]		This work	
	No. of features	Accuracy	No. of features	Accuracy	No. of features	Accuracy
Case 1: [A], [E]	3	100	4	100	3	100
Case 2: [D], [E]	3	100	4	100	3	100
Case 3: [A], [D], [E]	3	100	4	100	3	100
Case 4: [(A, B)], [(C,D)], [E]	3	80	4	80	3	100
Case 5: [(A, B, C, D)], [E]	3	100	4	100	3	100
Case 6: [A], [B], [C], [D], [E]	–	–	4	100	3	94.28

eyes) set B and set A are combined, inter-ictal subjects set C and set D (epileptogenic zone and hippocampal zone) are combined. Lastly, in case 5, again the binary classification was performed. However, again with a slight difference that four classes namely normal (A and B), inter-ictal (C and D) are combined and classification was done between the combined set and ictal set E.

The accuracy as reported for binary classification for both the case 1 and case 2 is 100%. The same accuracy of 100% has been reported for ternary classification for case 3 and case 4. As can be observed from the table, this study has achieved better results than both the studies performed by Alam and Bhuiyan [19] as well as by Khan and Farooq [20]. The same accuracy of 100% has been reported in case 5 as well. Alongside this, the number of features used in this study is less than what has been reported by Khan and Farooq, still demonstrated equal or higher performance. In addition to that, this study uses the median, which has been shown in various literature to be more robust to outliers than mean which was reported by Khan and Farooq. To further validate this method, a five-level classification has also been performed and compared with results from Khan and Farooq. The accuracy came out to be lesser than what has been reported by Khan and Farooq yet is sufficiently high enough to be used in the real world. The degradation in accuracy was because of misclassification between the opened eyes (set A) and closed eyes task (set B).

Results in Table 1 describes a very good classification accuracy for almost all levels. This suggests that this method can be used for detecting a seizure from an EEG signal. However, for clinical purposes, only two-level classification is usually required and this method has been proven to be as accurate as the other two methods used in comparison and much more computationally effective than them. The computational advantage over the methods used in comparison comes from the fact that Alam et al. [19] used EMD for decomposition of the signal and Khan et al. used wavelet decomposition and both of the methods are known to show poor computational performance.

## 6 Conclusion

In this study, nonlinear features such as Lyapunov exponent and approximate entropy were used to extract the underlying information from a dynamic EEG. In addition to that, a novel feature, i.e., Gini's coefficient was proposed. This study has shown that the proposed features were successful in capturing the relevant distinguishing information. Particularly, this study proves the proposed assertion that Gini's coefficient can be used for extracting distinctive information from the EEG signals of. This is evident from the results. Gini's coefficient is usually used as a metric to measure sparsity; this can be further used to study the sparsity of a seizure affected EEG signal in the time domain.

The proposed method is computationally less expensive and has performed better than the other methods reported in this study. Furthermore, the use of median has added robustness to this method. However, it has been reported in various literature

that the performance of approximate entropy degrades with increment in the length of time series. Therefore, the use of this particular feature may degrade the performance of the system.

In addition to that, the data under consideration have already been preprocessed and cleaned. This is not the case with clinical data, which contains muscular artifacts, ocular artifacts, and noise as well. These unwanted artifacts might increase the number of false-positive cases and thereby degrading the performance of the algorithm. More validation is required in order for this method to be effective.

## References

1. Fisher, R. S. et al. (2005). Epileptic seizures and epilepsy: Definitions proposed by the international league against Epilepsy (ILAE). *International Bureau for Epilepsy (IBE)*. *Epilepsia*, *46*, 470–472.
2. Elger, C. E., et al. (2006). Seizure prediction: The long and winding road. *Brain*, *130*, 314–333.
3. Birbeck, G. L. (2010). Epilepsy care in developing countries: part I of II. *Epilepsy Currents*, *10*, 75–79.
4. Rolston, J. D. et al. (2011). Electrical stimulation for epilepsy: Experimental approaches. *Neurosurgery Clinics of North America*, 425–v.
5. Herman, S. T. et al. (2015). Society, C.C.C.E.E.G.T.F. of the A.C.N.: Consensus statement on continuous EEG in critically ill adults and children, part I: indications. *Journal of Clinical Neurophysiology*, *32*, 87–95.
6. Wilson, S. B., et al. (2003). Seizure detection: Correlation of human experts. *Clinical Neurophysiology*, *114*, 2156–2164.
7. Tawfik, N. et al. (2015). A hybrid automated detection of epileptic seizures in EEG records.
8. Wang, L. et al. (2017). Automatic epileptic seizure detection in EEG signals using multi-domain feature extraction and nonlinear analysis.
9. Tzallas, A. T., et al. (2009). Epileptic seizure detection in EEGs using time-frequency analysis. *IEEE Transactions on Information Technology in Biomedicine*, *13*, 703–710.
10. Gajic, D., et al. (2015). Detection of epileptiform activity in EEG signals based on time-frequency and non-linear analysis. *Frontiers in Computational Neuroscience*, *9*, 38.
11. Andrzejak, R. G. et al. (2001). Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state, *64*, 1–8.
12. Mohseni, H. R. et al. (2006). Seizure detection in EEG signals: A comparison of different approaches, 6724–6727.
13. Kannathal, N. et al. (2005). Characterization of EEG—A comparative study.
14. Wolf, A. et al., Determining Lyapunov exponents from a time series alan WOLF. Jack B. SWIFT, Harry L. SWINNEY and John A. VASTANO, 285–317.
15. Kong, H. W. *20*, (1985), 2091–2094.
16. Hurley, N. et al. (2009). Comparing measures of Sparsity. *55*, 4723–4741.
17. Sen, A. (1973). *On economic inequality*. Clarendon Press. Oxford.
18. Haeb, A., & Ney, H., Linear discriminant analysis.
19. Alam, S. M., & Bhuiyan, M. (2013). Detection of seizure and epilepsy using higher order statistics in the EMD domain.
20. Khan, Y., Farooq, O. (2015). Wavelet-based multi-class discrimination of EEG for seizure detection. *19*, 266–278.

# Addressing Security and Privacy Issues of Load Balancing Using Hybrid Algorithm



T. Subha

**Abstract** In today's world, the need and urge for use of cloud become more popular among the public users. The cloud provides services like freeware to the end-users. The resources that the cloud users use will be in form of shared pool. If any resources are requested by the end-users, they are provided in a shared pool. Nowadays, the resources are requested only in dynamic basis. Upon the requisition by the user, the resources are provided to them. From these shared pools of resources, the cluster head or master node is selected by using Advanced Ant Colony optimization algorithm. The status of each and individual nodes should be known to neighbor nodes and master nodes; these can be achieved by using "Heartbeat messages". The status and movement of an individual node can be known by using these messages. The services requested by end-user and they are provided to them in very secure manner using DMZ (De-militarized zone) technique. The DMZ provides very higher security, that is, three layers of security, with different algorithms at each layer. In this paper, we address data leakage security issues and dynamic load balancing issues.

**Keywords** Ant colony · Cloud · De-militarized zone · Data leakage · Heartbeat messages · Hybrid algorithms · Load balancing · Optimization

## 1 Introduction

Nowadays everything has been changed to the internet. Cloud is a technology was evolving today to fulfill the user's need. They provide services to user's using internet. Shared resources are provided to end-user upon their individual requisition of that particular resource available in a pool [1]. The most prominent characteristics of cloud involve they want to store huge amount of data, so their storage capacity should be higher. We can extend the computing environment as dynamic basis, so elasticity and adapt to dynamic change in environment play a vital role here. The user can request any amount of data. Four types of infrastructure that cloud environment provides

---

T. Subha (✉)  
Sri Sai Ram Engineering College, Chennai, India  
e-mail: [subharajan@gmail.com](mailto:subharajan@gmail.com)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_16](https://doi.org/10.1007/978-981-15-0694-9_16)

based upon user's usage of services, they are: 1. Private cloud infrastructure (owned and used privately), 2. Public cloud infrastructure (publicly accessible), 3. Community cloud infrastructure (community-based group), 4. Hybrid cloud infrastructure (combination of public and private cloud) [2]. Cloud offers three types of services such as Software as a Service—SaaS, Infrastructure as a Service—IaaS, Platform as a Service—PaaS [3]. SaaS examples are Google online office, Google docs, Gmail, E-mail cloud, etc. In IaaS resources such as network, servers, software are provided to customers based on demand by the cloud service provider. Google App Engine is an example for PaaS [4]. This kind of provisioning of resources reduces capital investment and operational costs for industries and individual customers yielding better performance.

## 2 Load Balancing in Cloud

### 2.1 What Is Load Balancing?

Since the users of the cloud are huge in numbers. They store vast and wide range of data. For processing the data and storing the data, the cloud environment needs higher storage space [5, 6]. To attain higher performance, loads of individual nodes get shared among all nodes in particular cloud environment. The sharing of loads makes the nodes more efficient and retrieves the result faster. They follow certain algorithms to attain these balancing of loads in cloud environment [7]. It also ensures whether all the nodes or processors are sharing the load approximately at any point in time [4]. It provides solution for various issues in cloud computing. Load balancing is majorly categorized into provisioning of resources allocation and scheduling of tasks in distributed environment [8].

### 2.2 Objectives of Load Balancing

- Availability of resources based on demand.
- Efficient utilization of resources in spite of heavy or light load.
- Energy-saving under the circumstances of low load if the usage value of resources falls below the threshold.
- Cost optimization or minimization.

**Finally, customers expect a complete satisfaction, high efficiency in terms of provisioning a computing resource based on the best allocation strategies [9].**



## 2.3 Working of Load Balancing

It is the process of mapping the resources to different entities based on demand. Resource shall be allocated in such a way that no node must be overloaded and shall ensure no wastage of cloud resources in terms of memory, bandwidth processor, etc.

### 2.3.1 Resource Provisioning/Allocation

Mapping is being carried out in 2 levels as mentioned below.

**Mapping of the virtual machine to the host** Virtual Machine resides on the physical servers. Several Virtual machine instances are mapped on to the host based on its availability and processors capabilities. The provisioning policy determines how processing cores can be assigned to a virtual machine and it is responsible for the host that assigns Processing core to virtual machine [10] (Fig. 1).

**Mapping of application/ task to the virtual machine** Application/task requires processing power to complete their execution which takes place on the virtual machine. The virtual machine is responsible for providing the processing power to the tasks that are mapped on to it. It is done based on the configuration and availability.

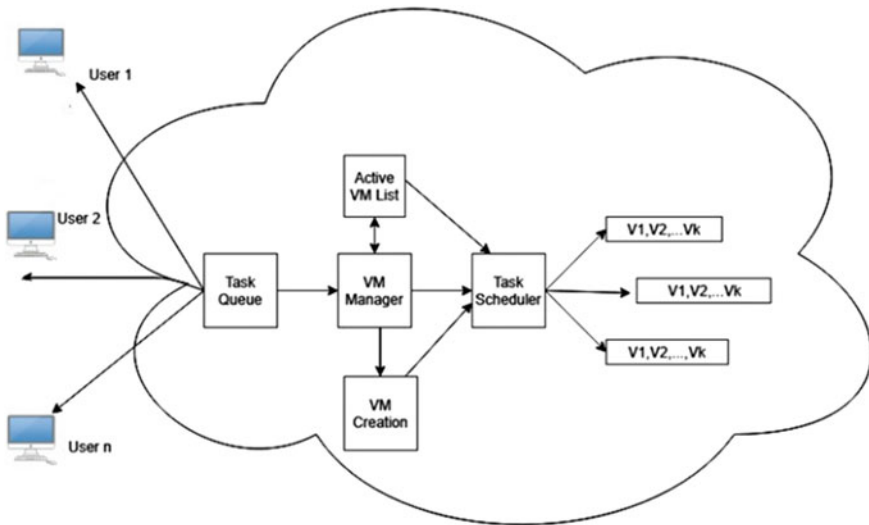


Fig. 1 Architecture of load balancing (referenced from [10])

### 2.3.2 Task Scheduling

This is the next level of activity after provisioning of resources in load balancing. It defines how the allocated resources are made available to the customers or end-users. It delegates the available resources fully until job completion (or) delegates the resources on a shared basis.

## 2.4 Issues in Load Balancing

There had been many security issues encountered in cloud storage. Security and privacy of the data stored in cloud storage have been addressed in [11]. There are several technical challenges such as server consolidation, security, fault tolerance, relocation of virtual machines, availability of resources, scalability, etc., have been encountered in load balancing. But, Load balancing is being identified as the central issue in cloud computing [12, 13].

**This chapter discusses load balancing and its various issues. It is identified that a system is required to address the security issue in load balancing.**

## 3 Algorithms for Load Balancing

Algorithms for sharing the load are majorly categorized into static algorithms for load balancing and dynamic algorithms for load balancing. The descriptions of all these algorithms are explained as follows:

### 3.1 Static Load Balancing Algorithm

All the nodes in the cloud environment are in still (static) state. The communication among all nodes is taken by means of passing messages. Every node in cloud environment will pass the message regarding the current status of node among the neighboring nodes and master nodes. These message transfers take place at initial time of communication. So at time of user communication, the current status of the node is not known.

Only the nodes previous state at the time of initialization is known and based on that, the communication takes place [12]. Tasks are assigned to processors in the compile-time itself before the execution of a program starts. These are non-preemptive scheduling algorithms and factors such as execution time of each task, requirement of resources should be known in prior. These algorithms are not suitable for systems that may change the load dynamically [14].

### **3.2 *Dynamic Load Balancing Algorithm***

In a changing load environment, all nodes are movable (dynamic). No node will be in a static state. All nodes initially register with the master node regarding their individual ID and maximum load they can manage. This communication takes place via messages. The messages are sending periodically to neighboring nodes and master nodes knowing the current status of individual nodes. By knowing the current status of individual node, it is easy to share the load among nodes [4]. The workload is shared either in a centralized or distributed manner in dynamic load balancing method.

Centralized algorithms are simple to implement and suffer from bottleneck problem and single point of failure [15]. The major drawback of this method is node starvation and resource allocation. Only the higher load balancing node will be given high priority and provide resources at the time. So low-priority node will not get the resources when they are required. This leads to Individual node starvation and resource allocation problem. Although this drawback, they are negligible, nowadays all organizations use dynamic load balancing algorithm.

### **3.3 *General Algorithms for Load Balancing***

This section briefly discusses the different types of load balancing algorithms that are used to share the load among nodes. In load balancing techniques, all the nodes transfer the load to nearest node by using certain algorithm. Let us see the overview presented in [16, 17]. The main advantage of this algorithm comes into the scenario of dynamic load environment the nodes in this environment share their load among all the nodes by knowing their status. In all cloud networks initially the master node or head can be elected using optimization algorithm. This master node knows the status of an individual node by periodically updating their individual status. The various load balancing algorithm as follows.

#### **3.3.1 *Ant Colony Optimization Algorithm***

This load balancing algorithm is based on ant colony optimization. It tries to balance the workload of the entire system by minimizing the makespan of the tasks. Many computational problems are solved using this probabilistic technique. Like the ant find the optimum path to find the foods, communication and data transfer took part in optimum way. This methodology has been integrated into CloudSim simulator toolkit. It outperforms FCFS and basic ant colony optimization algorithm [18].

### 3.3.2 Advanced Ant Colony Optimization Algorithm (AACO)

The nodes of the similar task are grouped together to form a batch. The nodes send the status along with the maximum load it can handle. These communications are taking place by “heartbeat messages”. They also generate UID to all nodes at session initialization. The threshold level of node is found using the node load. The higher capacity node can be elected as master node. If any priority node fails or misbehaves then the next priority node takes the responsibility of next higher priority nodes. The updation of every node status is initiated from every node to master node and find the optimum path to reach the master node [19, 20].

### 3.3.3 Round-Robin Algorithm

This scheduling algorithm selects the virtual machines randomly for placement. The controller in the data center does the assignment of VM's on a rotation basis. Initially the VM placement for first job request is chosen randomly. Then the VM for following requests has been assigned in a round-robin order, i.e., in circular order. The drawback of this scheme is execution time requirement of each process is not taken into account. The incoming job request needs to wait if the VM is not available for placement [21].

### 3.3.4 Throttled Load Balancing Algorithm

Here load balancer maintains the status of the VM (Available/Busy) in the index table. If any job request arrives, the data center sends a request to the load balancer for the VM availability. It scans the table until it finds the first available VM from the top. Then it communicates the VM id to the data center and updates the table. Further data center acknowledges this reply from load balancer. The load balancer sends -1 as a result if VM is not found suitably [22, 23].

### 3.3.5 Modified Throttled Algorithm

It works like throttled algorithm except that the index table is scanned from the first index that is previously assigned for the incoming job request for VM. The next VM present for the previously assigned VM is selected for placement. But this scheme is not always beneficial [24].

### 3.3.6 Min-Min Scheduling Algorithm

This algorithm begins with scanning all the tasks. It assigns the resource to the task that has the lowest completion time. Likewise, all the tasks are assigned resources

based on the minimum completion execution time. The existing load of a resource is not considered for allocation before a resource is assigned to a job. So it does not achieve proper load balancing [25].

### **3.3.7 Min-Min Scheduling Algorithm with Load Balancing**

This is a variation of Min-Min scheduling and uses this as a base for VM allocation. It consists of a request manager, service manager. Request manager receives a request for a task. Then request manager assigns it to a second-level service manager. Service manager divides the assigned task into subtasks and assigns it to processor for execution by taking the following factors such as CPU availability, transmission rate and memory into account [26]. But it is not suitable to assign large computation tasks.

### **3.3.8 Load Balance and User Priority Aware Improved Min-Min Scheduling Algorithm**

This improved algorithm initially starts execution by min-min algorithm. Next, it finds the tasks with minimum execution time from the heavily loaded resource. Then it checks with the makespan. If the value is less, it redistributes the task to the one that produces it. Hence the overloaded resources are freed. Idle or under loaded resources are utilized. A modification to this algorithm named user priority aware load balancing is proposed. In this, it assigns the resource based on the priority [25, 27].

### **3.3.9 Co-Operative Scheduling Anti Load Balancing Algorithm**

There are many algorithms proposed for load balancing in the cloud. This model considers the response time of jobs as the only criteria to compute the node capability. It takes weight of the node, threshold, overloaded resources and migration cost into account while assigning VM. The task with highest load will stay long on the host. It achieves better performance [8].

## **4 Cost-Effective Load Balancing Algorithm**

The following section describes the various cost-effective load balancing algorithms for cloud computing.

### ***4.1 Optimal Cost Scheduling Algorithm***

Round-robin algorithm is used to schedule the task based on cost. This, in turn, optimizes the cost [28]. In this, the resources are grouped as packages and placed inside a VM. When the resource is requested the VM that has the package is allotted. This tries to minimize the cost of execution at the service provider.

### ***4.2 Power-Aware Load Balancing Cost Scheduling Algorithm***

Every active compute node has the utilization parameter based on the service. It calculates the utilization value (i.e., percentage) of all active nodes in a network. If the percentage goes above 75%, a new VM has been instantiated and assigned a very low utilization percentage. Otherwise, if it is able to adapt a VM size, new VM is booted on a computer [29].

### ***4.3 Estimated Task Finish Time Cost Scheduling Algorithm***

This algorithm performs load balancing by estimating the finish time of the task before the job allocation. In this method, both the current load of the VM plus the time taken to finish the execution of a task is considered. This overcomes the problem of static load balancing algorithms [30].

### ***4.4 Optimal Peak Hour Performance in Data Centers Algorithm***

Peak hour performance load balancing algorithm verifies that the loads are equally distributed among all the nodes or processors during peak hours also. Because the requests received at peak hour is very high comparatively during normal hours. Hence faster response time has been ensured in this algorithm [31].

### ***4.5 Power Consumption Management Scheduling Algorithm (Bee-MMT)***

Artificial bee colony algorithm and Bee-MMT algorithm can be used to find overutilized nodes. It switches back and forth between the overutilized and underutilized

host based on the decisions taken by MMT scheduling algorithm [32]. It makes the underutilized node state to sleep after migrating all the tasks to another host.

#### ***4.6 Particle Swarm Optimization Task Based System Scheduling Algorithm***

This load balancing algorithm uses particle swarm optimization (PSO) algorithm. It identifies the extra task on an overloaded VM and transfers it into new VM. It does not migrate the entire workload rather only the extra tasks [33]. The pausing of VM is also avoided in the case of migration for a heavily loaded VM. It improves performance and reduces the time taken for processing compared with traditional load balancing approaches. Also the customer's last activity is presumed instead of lost from VM. It achieves customer satisfaction and improves QoS.

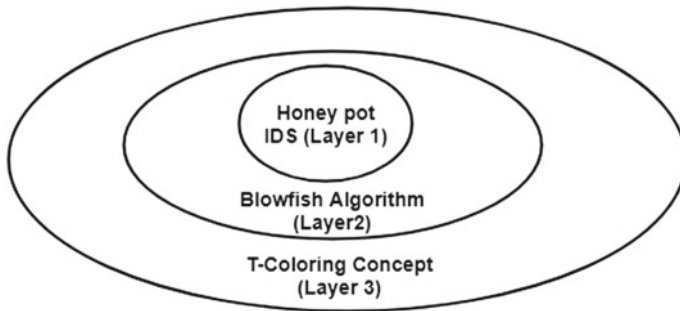
#### ***4.7 Cluster-Based Load Balancing Scheduling Algorithm***

The computing nodes are grouped into clusters in cluster-based load balancing algorithm. It comprises inter-cluster communication node (ICC), master and slave node. All the computing elements are termed slave nodes. All the slave nodes are connected to one master node. They update the recent values of following details like storage, processing capability, and bandwidth to their master node periodically [34]. All these details are maintained in a table by master node. An individual entry in a table represents a load of each and every slave nodes in a network. The same load balancing distribution is carried out in two methods [35, 36]. In first method, load is shared within the master nodes. The decision is taken based on the calculation of performance factor. In the second part, the load is distributed from master node to slave node based on round-robin algorithm. Finally, the algorithm achieves better execution time, waiting time, turnaround time, and high throughput [37].

**This chapter helps to understand the different available load balancing algorithms and QoS parameters to be improved. It also specifies the advantages and disadvantages of every algorithm. It provides a way to design a new security algorithm for solving security issues in load balancing.**

### **5 Proposed Hybrid Load Balancing and Security Algorithm**

We propose our integrated load and security algorithm in this section. This algorithm aims in finding the most secure area of transmission in the cloud and provide optimum



**Fig. 2** Proposed Architecture of load balancer for secure data transmission

load balancing scheme. They provide the integration of both load balancer and secure data transmission. The load balancing technique is taken care by intermediate router and secure transmission is achieved by means of DMZ (De-militarized zone). The DMZ forms a three-layer of security from the client node to server node. Therefore, upon user requisition the secure data transfer is provided to them.

So far we had seen regarding various load balancing algorithms, Layer 1 deals with Honey pot IDS, layer 2 deals with the blowfish algorithm, and layer 3 deals with T coloring concept for segregation of nodes. Honey pot act as an intrusion detection system, they check the malicious traffic and alert the system administrator. Thereby they get alert from future attack. The honey pot is a bribed component for promising nodes. They compromise themselves and behave in favor of administrator for IDS.

Layer 2 deals with blowfish security algorithm, blowfish deals with an encryption algorithm and most efficient and flexible one. It works under two methods, namely, key expansion and data encryption (Fig. 2).

Layer 3 deals with T coloring concept in which the nodes get segregated and divide the files into various parts, so that intruders find difficult to trace and locate. T coloring concept introduces splitting the files among the nodes. The nodes are placed under certain conditions no two nodes are in an adjacent direction.

The proposed methodology of hybrid load balancing with a security algorithm can be implemented anywhere in parallel and distributed clouds. It is purely applicable to the areas such as health care, resource allocation in cloud, storing data in cloud, analytics, etc., where enormous amount of sensitive and private data are generated. Hence it requires the guarantee of data usage in a proper way i.e. it should not be tampered or altered.

## 6 Conclusion

As cloud computing plays an important role in today's world, we have illustrated the advantages of load balancing and various types of load balancing algorithms applicable to homogeneous and heterogeneous environment in detail. The comparison of



different static and dynamic load balancing algorithms is listed in table. This paper mainly addresses the security issues of load balancing. A methodology has been proposed to combine load balancing concepts with security algorithm to overcome this. Our proposed approach achieves optimum load balancing by finding a secure area for transmission. It is achieved via DMZ and a three-layer security mechanism. So, our system can attain higher performance throughput and less fault tolerance rate by integrating various IDS and IPS rule. Further, data leakages in cloud are also avoided by this technique. In future work, we propose to build various three-layer security mechanism to attain higher security rates in cloud domain in real time.

## References

1. Pantazoglou, Michael, Tzortzakis, Gavril, & Delis, Alex. (2016). Decentralised and energy-efficient workload management in enterprise clouds. *IEEE Transactions on Cloud Computing*, 4(2), 196–209.
2. Ran, Chen, Wang, Shaowei, & Wang, Chonggang. (2015). Balancing backhaul load in heterogeneous cloud radio access networks. *IEEE Wireless Communication*, 22(3), 42–48.
3. Zhao, Jia, Yang, Kun, Wei, Xiaohui, et al. (2016). A heuristic clustering-based task deployment approach for load balancing using Bayes theorem in cloud environment. *IEEE Transactions on Parallel and Distributed Systems*, 27(2), 305–316.
4. Nahir, Amir, Orda, Ariel, & Raz, Danny. (2016). Replication-based Load balancing. *IEEE Transactions on Parallel and Distributed Systems*, 27(2), 494–507.
5. Tianyi Chen, Yu., Zhang, Xin Wang, & Giannakis, Georgios B. (2016). Robust Workload and Energy management for sustainable data centers. *IEEE Journal on Selected Areas in Communications*, 34(3), 1.
6. Octavio Gultierrez-Garcia, J., & Nafarrate, Adrian Ramirez-. (2015). Collaborative agents for distributed load management in cloud data centers using live migration of virtual machines. *IEEE Transaction on Services Computing*, 8(6), 916–929.
7. Xiaolong, Xu, Cao, Lingling, & Wang, Xinheng. (2016). Adaptive task scheduling strategy based on dynamic workload adjustment for heterogeneous Hadoop clusters. *IEEE System Journal*, 10(2), 471–482.
8. Beloglazov, A., & Buyya, R. (2010). Energy efficient resource management in virtualized cloud data centers. Paper presented at the 10th IEEE/ACM international conference on cluster, cloud and grid computing, pp. 826–831.
9. Evers, X., CSG, W. H., CR.B.SG. (1992). A literature study on scheduling in distributed systems, Delft university Of Technology.
10. Singh, Athokpam B., Sathyendra Bhat, J., Raju, Ragesh, & D’Souza, Rio. (2017). Survey on various load balancing techniques in cloud computing. *Advances in Computing*, 7(2), 28–34.
11. Subha, T., & Jayashri, S. (2017). Public auditing scheme for data storage security in cloud computing. *Journal of Information Science and Engineering*, 33, 773–787.
12. Qi Zhang, Lu, & Cheng, Raouf Boutaba. (2010). Cloud computing: state-of-the-art and research challenges. *Journal of Internet Services and Applications*, 1(1), 7–18.
13. Caron, E., Rodero-Merino, L., Desprez, F., & Muresan, A. (2012). Auto-scaling, load balancing and monitoring in commercial and open-source clouds, Research Report no. 7857.
14. Mishra, R., & Jaiswal, A. (2012). Ant colony optimization: a solution of load balancing in cloud. *International Journal of Web & Semantic Technology*, 3(2), 33.
15. Sidhu, Amandeep Kaur, & Kinger, Supriya. (2013). Analysis of load balancing techniques in cloud computing. *International Journal of Computers & Technology*, 4(2), 471–737.

16. Nuaimi, K. A., Mohamed, N., Nuami, M. A., Al-Jaroodi, J. (2012). A survey of load balancing in cloud computing: challenges and algorithms. Paper presented at the 2012 Second Symposium on Network Cloud Computing and Applications (NCCA), pp. 137–142.
17. Wang, S.-C., Yan, K.-Q., Liao, W.-P., & Wang, S.-S. (2010). Towards a load balancing in a three-level cloud computing network. In *Proceedings of the 3rd International Conference on Computer Science and Information Technology (ICCSIT)*, pp. 108–113.
18. Fahim, Y., Ben Lahmar, E., Labriji, E. H., & Eddaoui, A. (2014). The load balancing based on the estimated finish time of tasks in cloud computing. *Proceedings of the Second World Conference on Complex Systems (WCCS), 2014*, 594–598.
19. Kuhl (1998) A Taxonomy of Scheduling in General-purpose Distributed Computing Systems, *IEEE transactions on software engineering* 14(2):141–154.
20. Shah, M. M. D., Kariyani, M. A. A., Agarwal, M. D. L. (2013). Allocation of virtual machines in cloud computing using load balancing algorithm. *IJCSITS* ISSN: 2249-9555.
21. Wickremasinghe, B. (2009). CloudAnalyst: a cloudSim-based tool for modeling and analysis of large scale cloud computing environments. *MEDC Project Report*, 22(6), 433–659.
22. Wickermasinghe, B., Calheiros, R. N., & Buyya, R. (2010). Cloudanalyst: a cloudsim-based visual modeller for analysing cloud computing environments and applications. Paper presented at the 24th IEEE International conference on advanced Information Networking and Applications (AINA), pp. 446–452.
23. S.G. Domanal, G.R.M. Reddy (2013) Load Balancing in Cloud Computing using Modified Throttled Algorithm. Paper presented at the IEEE International Conference on Cloud Computing in emerging Markets(CCEM), pp. 1–5.
24. Elian, G. A. (2013). User-priority guided Min-Min Scheduling algorithm for load balancing in cloud computing. Paper presented at the National Conference on Parallel Computing Technologies (PARCOMPTECH), pp. 1–8.
25. Thiam, C., Da Costa, G., & Pierson, J. M. (2013). cooperative scheduling anti-load balancing algorithm for cloud: CSAAC, Paper presented at the 5th International Conference on Cloud Computing Technology and Science (CloudCom), pp. 433–438.
26. Yao, J., & He, J. H. (2012). Load balancing strategy of cloud computing based on artificial bee algorithm. Paper presented at the 8th International Conference on Computing Technology and Information Management (ICCM), pp. 185–189.
27. Galloway, J. M., Smith, K. L., & Vrbsky, S. S. (2011). Power aware load balancing for cloud computing. *Proceedings of the World Congress on Engineering and Computer Science, 2011*, 19–21.
28. Soundarabai, P. B., Rani, A. S., Sahai, R. K., Thriveni, J., Venugopal, K. R., & Patnalk, L. M. (2014). Load balancing with availability checker and load reporters (LB-ACLRs) for improved performance in distributed systems. In *Proceedings of the 2nd International Conference on Devices, Circuits and Systems (ICDCS)*, pp. 1–5.
29. Chawla, A., & Ghumman, N. S. (2015). Efficient cost scheduling algorithm with load balancing in a cloud computing environment. *International Journal of Innovative Research in Computer and Communication Engineering*, 3(6).
30. Sumalatha, M. R., Selvakumar, C., et al. (2014). CLBC-cost effective load balanced resource allocation for partitioned cloud system. *Proceedings of the International Conference on Recent Trends in Information Technology (ICRTIT), 2014*, 1–5.
31. Achar, R., Thilagam, P. S., Soans, N., Vikyath, P. V., Rao, S., & Vijeth, A. M. (2013). Load balancing in cloud based on live migration of virtual machines. *Proceedings of the Annual IEEE India Conference (INDICON), 2013*, 1–5.
32. Kulkarni, A. K., & Annappa, B. (2015). Load balancing strategy for optimal peak hour performance in cloud datacenters. In *International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES)*, pp. 1–5.
33. Li, K., Xu, G., Zhao, G., Dong, Y., et al. (2011). Cloud task scheduling based on load balancing ant colony optimization. *Proceedings of the Sixth Annual China grid Conference (ChinaGrid), 2011*, 3–9.

34. Ramezani, F., Lu, J., & Hussain, F. K. (2014). Task-based system load balancing in cloud computing using particle swarm optimization. *International Journal of Parallel Programming*, 42(5), 739–754.
35. Dhurandher, S. K., Obaidat, M. S., Woungang, I., et al. (2014). A cluster-based load balancing algorithm in cloud computing. *Proceedings of the IEEE International Conference on Communications (ICC), 2014*, 2921–2925.
36. Kapoor, S., & Dabas, C. (2015). Cluster based load balancing in cloud computing, In *Proceedings of the Eighth International Conference in Contemporary Computing (IC3)*, pp. 76–81.
37. Katyal, Mayanka, & Mishra, Atul. (2013). A comparative study of load balancing algorithms in cloud computing environment. *International Journal of Distributed and Cloud Computing*, 1(2), 1–13.

# Key Management Scheme for Secure Group Communication



Om Pal and Bashir Alam

**Abstract** Multicast or group communication enables the distribution of the content in a one-to-many fashion. In multicast communication, the major challenges are dynamicity of group, forward and backward secrecy of the data. There are issues like single-point failure in centralized Group Key Management (GKM), false participation attack in participatory GKM, member dynamicity in Logical Key Hierarchy, etc. To address the various issues of centralized, participatory and LKH GKM; in this paper, we proposed a Key Management Scheme for Secure Group Communication. In the proposed scheme, network members share the computational load of the server and scheme achieves the forward and backward secrecy. The proposed scheme is well suitable for one-to-many mode communication.

**Keywords** Multicast communication · Group key management · Key distribution · Key management

## 1 Introduction

In present days, most of the cybersecurity applications use the Group Key (GK) for encrypting and decrypting the common information. In such systems, Group Controller (GC) or any member of the group multicasts the common information to other members or sub-members of the group. Multimedia transmission, distance learning, video conference, data replication, defense systems, distributed network, cloud computing, multi-party games, etc., are some areas where common information is transmitted in one-to-many mode. In multicast communication, due to single encryption of common information, bandwidth is saved, timely delivery of data is ensured, quality of service is improved.

---

O. Pal (✉)

Ministry of Electronics and Information Technology Government of India, New Delhi, India  
e-mail: [ompal.cdac@gmail.com](mailto:ompal.cdac@gmail.com)

O. Pal · B. Alam

Faculty of Engineering & Technology, Department of Computer Engineering, Jamia Millia Islamia, New Delhi, India

With various benefits of multicast communication, there are also many challenges which include communication overhead, scalability, forward secrecy, backward secrecy, storage cost, key initialization cost, computational overhead, etc. In dynamic multicast communication, members join and leave the group frequently and due to frequent leave and join request, maintaining the forward and backward secrecy of the group are major concerns for any Key Server. To achieve the forward secrecy, it is ensured that leaving member should not be able to decipher the future messages of the group. In backward secrecy, it is ensured that new member should not be able to decipher the past messages.

Authors have presented various Group Key Management (GKM) Schemes in the literature which includes Logical Key Hierarchical structure based schemes, centralized, de-centralized, contributory and participatory schemes. To overcome the limitations of the Logical Key Hierarchy (LKH) architecture, VijayKumar et al. [1] proposed a centralized group key distribution scheme. In this scheme, some computational parameters are multi-casted and using these received parameters, existing members of the group, derive the updated group key. This scheme suffers with problem of forward secrecy. Leaving member is able to read the future messages if the key ( $K_1$ ) of leaving member, completely divides the key ( $K_i$ ) of any existing member of the group. Another drawback of this scheme is that Group Controller suffers with high computational overhead during the updation of existing key. During the updation of existing key, it is ensured that all keys ( $K_s$ ) should be greater than  $\mu$ .

To overcome the limitations of the scheme [1], in this paper we proposed the key management for multicast communication. To minimize the fail attempts, we used two multiplicative algebraic groups. We limit the range of  $\psi$  by selecting  $\Theta \in z_{pr_1}^*$  and restricting  $q < pr_2$ . Range of  $sk_i$  is higher than both parameters  $\psi$  and  $\Theta$  due to selection of  $sk_i$  from  $z_{pr_2}^*$ . Due to the above advantage, we applied the ceiling  $pr_1 \leq \lceil \Gamma pr_2/4 \rceil$ ,  $pr_2 > \zeta$  and  $\zeta \leq \lceil \Gamma pr_2/4 \rceil$ . Due to this change, unnecessary computation at GC side is avoided.

In scheme [1] leaving member can compute the inverse of 'a' and using it, leaving member can decrypt the future messages of the group. To maintain the forward secrecy, it is necessary that  $sk_i$  must not be a factor of any of the other keys. In our proposed scheme, we eliminated this attack.

## 2 Related Work

Many solutions have been proposed for key management in secure group communication [1–22]. VijayKumar et al. [1] have presented a solution for key distribution using extended Euclid algorithm. The authors have significantly reduced computational complexity. However, the forward secrecy is not maintained in the protocol. VijayKumar et al. [2] proposed a Chinese remainder theorem based group key management scheme. The computational complexity of scheme is reduced up to  $O(1)$ . Storage complexity is also reduced. However, the group initialization requires very high computations. Based on RSA cryptosystem Kumar et al. [5] proposed a key

management protocol. The authors have reduced the computation, communication and storage complexities. Moreover, the protocol is secure against various attacks.

Iuon-Chang Lin et al. [23] proposed a multicast communication scheme based on RSA public cryptosystem. The proposed solution does not require rekeying after any changes in group membership. However, the protocol is not scalable for large groups. Kumar et al. [8] have proposed a non-interactive key agreement protocol for group communication. The authors reduced computation complexity. Using Chinese remainder theorem, Kumar et al. [17] have proposed a RSA based public-key cryptographic protocol. The protocol is secure against the factorization attack. Sharma et al. [18] have presented distributed key management scheme using elliptic curve cryptography. The protocol reduced computational complexity up to some level. However, mass join and mass leave operations not supported by the protocol.

Pal et al. [16] presented the one-to-many mode communication scheme for conditional access system. In this scheme, only Group Controller is able to pass the Control Word (CW) to the members of the channel. However, there is no one-to-many mode communication among the members of the group. In the real world, there are many applications where one-to-many mode communication is desirable for each member of the group. In this paper, the authors extended the scheme [16] and using the extended scheme, members of the group can communicate with each other securely. Any member of the group can send the common message to other members of the group in one-to-many mode.

In centralized GKM, most of the key computation works are done by the server. So, there are the issues in centralized GKM like single-point failure, the computational delay is proportional to number of existing members, etc. In participatory GKM, there are chances of man-in-middle attack, false participation attack, delay in Group Key computation, etc. In LKH architecture, more computational overhead and delay are involved if common data is sent to the group in which members belong from various sub-group of the LKH architecture. To overcome the various limitation of GKM, VijayKumar et al. [1] presented the group key distribution scheme. However, scheme [1] suffers with problem of forward secrecy. To address the various issues of centralized, participatory and LKH GKM; in this paper, a Key Management Scheme for secure group communication is proposed.

### 3 Proposed Group Communication Scheme

Let there be a Group Controller (GC), which prepares the setup for the network. To initialize a setup for 'n' members, Group Controller selects a number  $\zeta$  and two prime numbers  $pr_1, pr_2$  whereas  $pr_1 \leq \Gamma pr_2/4$ ,  $pr_2 > \zeta$  and  $\zeta \leq \Gamma pr_2/4$ . Over primes  $pr_1, pr_2$ , the algebraic groups  $Z_{pr_1}^*$  and  $Z_{pr_2}^*$  are formed. GC selects a random element  $\alpha$  from  $Z_{pr_2}^*$ .

### 3.1 Member Join

Using the public key of new user  $U_i$ , GC sends the random key  $sk_i \in Z_{pr_2}^*$  to user  $U_i$  for joining the network. GC reselects the  $sk_i$  if  $sk_i < \psi$  or LCM of secret keys of existing users is completely divisible by the secret key of the joining member.

Using the following steps, GC computes the Group Key (GK) and distributes GK to new and existing members of the network.

1. GC selects  $\Theta \in Z_{pr_1}^*$  randomly and computes the threshold parameter  $\psi = \Theta + \zeta$ . Here  $\psi$  is directly proportional to  $\Theta$ .
2.  $GK = \alpha^\Theta \pmod{pr_2}$ .
3. GC computes the value  $\lambda$  using the multiple (prod) of existing secret keys and secret key of new member:  $\lambda = \text{prod} \times sk_i$ . Initial value of prod is 1.
4. Using the extended Euclidian algorithm [21], GC derives the value of a:

$$a \times \Psi + b \times \lambda = 1 \quad (1)$$

5. GC multicasts values  $\alpha$ , a,  $pr_2$  and  $\zeta$ .
6. Any existing or new user computes the GK using the following steps (Let  $U_i$  computes GK)

- (i)  $\psi = a^{-1} \pmod{sk_i}$
- (ii)  $\Theta = \psi - \zeta$
- (iii) Finally  $U_i$  computes:  $GK = \alpha^\Theta \pmod{pr_2}$ .

### 3.2 Member Leave

There is a need to maintain the forward secrecy whenever any member leaves the network. To maintain the forward secrecy, GC recomputes the GK and distributes it to the remaining members of the network. Let leaving user is  $U_i$ . GC takes a new  $\Theta \in Z_{pr_1}^*$  randomly and computes a new threshold parameter  $\psi = \Theta + \zeta$ . GC updates the GK using the following steps:

1.  $GK = \alpha^\Theta \pmod{pr_2}$ .
2. New value of  $\lambda = \text{prod} / sk_i$ .
3. Using the extended Euclidian algorithm [21], GC derives the value of a:

$$a \times \psi + b \times \lambda = 1 \quad (2)$$

4. GC multicasts parameter 'a' to remaining members.
5. Remaining members computes GK using following steps (let remaining user  $U_r$  computes GK)

- (i)  $\psi = a^{-1} \bmod sk_r$
- (ii)  $\Theta = \psi \rightarrow \zeta$
- (iii) Finally  $U_r$  computes:  $GK = \alpha^\Theta \bmod pr_2$ .

## 4 Security Analysis

Users derive the GK using its secret key and other parameters sent by GC through multicast. For joining phase, GC multicasts  $\alpha, a, pr_2$  and  $\zeta$ . The intruder may capture the multicast parameters and he/she may try to deduce the GK using the received parameters. To obtain GK, confidential parameter  $\Theta$  is required so without having  $\Theta$ , intruder cannot obtain GK. Now intruder may try to deduce  $\Theta$  but for this parameter  $\psi$  is needed. To obtain the parameter  $\psi$  from transmitted value 'a', there is a need for secret key of any existing member. So, intruder can derive GK only if he/she has secret key of any existing member of the network. Without having the secret key, intruder cannot obtain the GK.

Proposed scheme achieves the forward secrecy. In case of any member leave from the network, GC excludes the secret key of leaving member from the database and it updates the value of  $\lambda$ . Due to new value of  $\lambda$ , leaving member cannot get the modified value of  $\psi$  therefore, intruder cannot compute the  $\Theta$  which is mandatory to compute the GK.

Proposed scheme also achieves the backward secrecy. Whenever any new member joins the network, GC updates the value of GK and distributes the updated GK to the new and existing members of the network. The new member does not have the old GK therefore; new member cannot decrypt the past messages of the network.

To obtain  $\Theta$  from  $z_{pr1}^*$ , the intruder may try to apply brute force attack but it is not feasible to obtain  $\Theta$  from  $z_{pr1}^*$  due to large size of the algebraic group  $z_{pr1}^*$ . Let size of the  $\Theta$  be 128 bits then total trials would be  $2^{127}$ . Let one trial be completed in 1  $\mu$ s then average time to guess  $\Theta$  would be around  $4.46 \times 2^{80}$  years. Therefore, it is concluded that guessing of GK is not feasible through brute force attack.

## 5 Applications and Future Scope

Scheme is useful in various fields like multimedia transmission, distance learning, video conference, data replication, defense systems, distributed network, cloud computing, multi-party games, etc. Against the one-to-one encryption, common data can be sent in one go to the group members. Due to single encryption of data, bandwidth may be saved, delay in delivery of common data may be reduced.

Group Key is a special kind of symmetric key. Group Key can be used for encryption of common data. Proposed Group Key Management scheme can be incorporated in emerging technologies like Blockchain for securing the sensitive records for the consensus mechanism. The scheme is also useful for IoT systems for sending the data in one-to-many communication mode.



## 6 Conclusion

In this paper, we analyzed the group key management schemes, major challenges of the key management schemes like forward and backward secrecy. We analyzed the scheme proposed by VijayaKumar et al. [1] and to remove the limitation of scheme [1], we proposed a new key management scheme. Security analysis of the proposed scheme is also done and we concluded that our proposed scheme achieves the forward secrecy and backward secrecy of the data and it reduces the computational load of the server. Scheme is well suitable for one-to-many mode communications.

**Acknowledgements** This publication is an outcome of the R&D work undertaken by a project under the Visvesvaraya Ph.D. Scheme of Ministry of Electronics & IT, Government of India, being implemented by Digital India Corporation.

## References

1. VijayaKumar, P. (2013). Centralized key distribution protocol using the greatest common divisor method. *Computers & Mathematics with Applications*, 65(9), 1360–1368.
2. VijayaKumar, P., Bose, S., & Kannan, A. (2014). Chinese remainder theorem based centralized group key management for secure multicast communication. *IET information Security*, 8(3).
3. Wallner, D. M., Harder, E. J., & Agee, R. C. (1998). Key management for multicast: issues and architectures. Internet Draft Report, Filename: draft-wallner-key-arch-01.txt.
4. Wallner, D., Harder, E., & Agee, R. (1999). Key Management for Multicast: Issues and Architectures. RFC 2627.
5. Kumar, V., Kumar, R., & Pandey, S. K. (2018). A computationally efficient centralized group key distribution protocol for secure multicast communications based upon RSA public key cryptosystem. *Journal of King Saud University Information Sciences*. <https://doi.org/10.1016/j.jksuci.2017.12.014>.
6. Steiner, M., Tsudik, G., & Waidne, M., Diffie-hellman key distribution extended to group communication. IBM Ziirich Research Laboratory CH-8803 Riischlikon, Switzerland.
7. Liu, Z., Lai, Y., Ren, X., & Bu, S. (2012). An efficient LKH tree balancing algorithm for group key management. In *2012 International Conference on Control Engineering and Communication Technology*, Liaoning, pp. 1003–1005.
8. Kumar, V., Kumar, R. & Pandey, S. K. (2018). Polynomial based non-interactive session key computation protocol for secure communication in dynamic groups. *International Journal of Information Technology*, pp. 1–6. <https://doi.org/10.1007/s41870-018-0140-1>.
9. Sherman, A. T., & Mcgrew, D. A. (2003). Key establishment in large dynamic groups using one-way function trees. *IEEE Transactions On Software Engineering*, 29(5).
10. Odelu, V., Das, A. K., & Goswami, A. (2016). A secure effective dynamic group passwordbased authenticated key agreement scheme for the integrated EPR information system. *Journal of King Saud University-Computer and Information Sciences*, 28(1), 68–81.
11. Xu, L., & Huang, C. (2008). Computation-efficient multicast key distribution. *IEEE Transactions on Parallel and Distributed Systems*, 19(5), 577–587.
12. Rafaeli, S., & Hutchison, D. (2002). Hydra: a decentralized group key management. In *11th IEEE International WETICE: Enterprise Security Workshop*, June 2002.
13. Hanatani, Y., et al. (2016). Secure multicast group management and key distribution in IEEE 802.21. Security Standardisation Research Springer International Publishing, pp. 227–243.

14. Ballardie, A. (1997). Core Based Trees (CBT version 2) Multicast Routing protocol specification", September 1997. RFC 2189.
15. Baddi, Y., & Dafir Ech-Cherif El Kettani, M. (2013). Key management for secure multicast communication: A survey. Security Days (JNS3), 2013 National. IEEE.
16. Pal, O., & Alam, B. (2019). Efficient and secure conditional access system for pay-TV systems. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-019-7257-5>.
17. Kumar V., Kumar R., & Pandey S. K. (2018) An Enhanced and Secured RSA Public Key Cryptosystem Algorithm Using Chinese Remainder Theorem, in third International Conference on Smart and Innovative Trends in Next Generation Computing Technologies (NGCT 2017). Communications in Computer and Information Science, vol 828. Springer, Singapore, pp. 543–554, [https://doi.org/10.1007/978-981-10-8660-1\\_42](https://doi.org/10.1007/978-981-10-8660-1_42).
18. Sharma, S., & Krishna, C. R. (2015). An efficient distributed group key management using hierarchical approach with elliptic curve cryptography. In *2015 IEEE International Conference on Computational Intelligence & Communication Technology, Ghaziabad*, pp. 687–693. <https://doi.org/10.1109/cict.2015.116>.
19. Amir, Y., Kim, Y., Rotaru, C. N., Schultz, J., Stanton, J., & Tsudik, G. (2019). Exploring Robustness in Group Key Agreement. Retrieved March 15, 2019, from <http://www.cnds.jhu.edu/pub/papers/cnds-2000-4.pdf>.
20. Adusumilli, P., & Zou, X. (2005). KTDCKM-SDC: a distributed conference key management scheme for secure dynamic conferencing. In *Proceedings of the 10th IEEE Symposium on Computers and Communications (ISCC 2005)*.
21. Naranjo, J. A. M., Lopez-Ramos, J. A., & Casado, L. G. (2010). Applications of the extended Euclidean algorithm to privacy and secure communications. In *Proceedings of the 10th International Conference on Computational and Mathematical Methods in Science and Engineering, CMMSE*.
22. Islam, Hafizul, S. K., & Biswas, G. P. (2017). A pairing-free identity-based two-party authenticated key agreement protocol for secure and efficient communication. *Journal of King Saud University-Computer and Information Sciences*, 29(1), 63–73.
23. Lin, I.-C., Tang, S.-S., & Wang, C.-M. (2010). Multicast key management without rekeying processes. *The Computer Journal*, 53(7), 939–950.

# Lightweight Hardware Architecture for Eight-Sided Fortress Cipher in FPGA



Nivedita Shrivastava and Bibhudendra Acharya

**Abstract** In the lightweight domain, various ciphers and their different implementations are introduced to deal with the problem of security in resource scarce environment. Eight-sided fortress (ESF) is a lightweight Feistel cipher which uses substitution–permutation network based round function with Serpent Substitution-box(S-box). This work presents a study and comparison of the various hardware architectures of ESF to combat issues of security in an extremely constrained resource environment. For the design of hardware, different techniques of S-box implementation are used. Comparison and evaluation of ESF S-box implementation techniques is done on the basis of latency, throughput, area utilization, and power consumption. It is observed that the Random Access Memory (RAM)-based S-box design gave the best results with the requirement of minimum area for its implementation. This makes it the preferred architecture for resource-limited applications.

**Keywords** ESF · BRAM · Synchronous · LUT · Feistel

## 1 Introduction

With the advancement in various tiny computing devices like radio frequency identification tag and sensor network nodes, there is a rise in demand for encryption techniques for these highly resource-constrained environment applications. Small embedded devices found their use in many of the applications [1]. For this purpose, various hardware architectures have been proposed which aim in providing an optimal trade-off between security, area requirement, and power consumption. This gives rise to the field of lightweight cryptography. Various existing and efficient cryptographic ciphers like data encryption standard (DES) [2] and advanced encryption

---

N. Shrivastava (✉) · B. Acharya  
Department of Electronics and Telecommunication Engineering, National Institute of Technology Raipur, Raipur, India  
e-mail: [nivedita.shrivastava01@gmail.com](mailto:nivedita.shrivastava01@gmail.com)

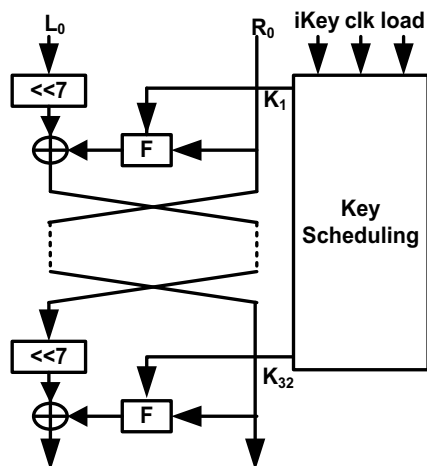
B. Acharya  
e-mail: [bacharya.etc@nitrr.ac.in](mailto:bacharya.etc@nitrr.ac.in)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_18](https://doi.org/10.1007/978-981-15-0694-9_18)

standard(AES) [3] are not apt for highly resource-limited applications [4]. Therefore, for this lightweight domain, various other algorithms have been proposed which are mainly categorized as block ciphers and stream ciphers. Some of the famous block ciphers are RECTANGLE [1], PRESENT [5], PRINT [6], HIGHT [7], LED [8], LBLOCK [9], KLEIN [10], and Eight-sided fortress (ESF) [11] while some of the famous stream ciphers are Grain [12] and Trivium [13] (Fig. 1).

Various implementation techniques are used for designing the hardware of these lightweight ciphers so as to achieve an optimum trade-off between security, resource consumption, privacy, and power consumption. Some of the implemented techniques are round-based architecture, iterative architecture, serial architecture, parallel architecture, pipelined, and many more. Ciphers are implemented in two main forms: Feistel network based and substitution—permutation network (SPN) based. Implementations based on Feistel ciphers generally provide slow diffusion which may lead to increased security concerns. So, to combat this problem, traditional Feistel ciphers use more number of encryption rounds as compared to ciphers based on SPN which requires lesser number of rounds [14]. There are certain merits of Feistel ciphers as well. The first one is that they use simple, small, and easy to implement round functions. The second one is that they generally require the same program for the implementation of both encryption and decryption part, thus reducing resource requirement and making it useful for various lightweight applications.

In this work, a new hardware implementation technique for S-box implementation in case of the ESF algorithm is proposed. Comparison and evaluation of S-box implementation technique is done in this paper on the basis of hardware requirements for its implementation in field programmable gate array (FPGA), power consumption, and latency of architecture. In this work, the target algorithm is ESF which is a Feistel network based cipher with SPN round function which uses Serpent S-boxes



**Fig. 1** Eight-sided fortress algorithm [11]

for performing the substitution. The round-based architecture of ESF cipher is taken into consideration for implementation as well as for the purpose of calculation and evaluation of results. Along with this technique, key retrieval methods based on memories are used and evaluated to get the optimum results for the implemented architectures.

This paper is organized into six sections. Section 2 summarizes basics of ESF algorithm and explains details of the round function, substitution layer, and permutation layer used in the implementation of ESF cipher. Section 3 describes the proposed hardware architecture of ESF cipher for round-based architecture with different S-box implementation techniques and key retrieval mechanism. Section 4 shows the results of the work along with a comparison with other works in the field. Section 5 is about evaluation and discussion of the results obtained for all the implemented architectures and finally, Sect. 6 is the conclusion of the work.

## 2 ESF: Algorithm

ESF is a Feistel network based cipher which uses SPN-based round function. ESF takes 64-bit plaintext as input and uses a key length of 80 bits. It performs 32 rounds of the encryption, as shown in Fig. 2. The encryption process of ESF for a single round can be summarized as given below:

- 64 bits of input plaintext of ESF is divided into two parts, leftmost 32 bits and rightmost 32 bits. Rightmost 32 bits are processed via round function whose description is given in the next section while leftmost 32 bits are shifted left by 7 bits.
- Thereafter, processed rightmost 32 bits are XORed with 7 bits shifted leftmost bits.
- In the end, swapping of the leftmost and rightmost bits is done.

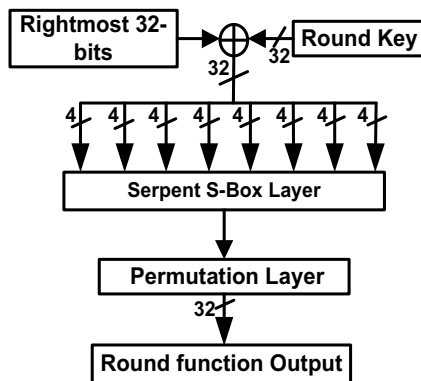


Fig. 2 ESF round function [11]

- And finally, after completion of all 32 rounds, the final ciphertext is obtained. The right part gives (Least significant bit) LSB side bits while the left part gives (Most significant bit) MSB side bits of the ciphertext.

## 2.1 Round Function

Rightmost 32 bits are processed via a round function which is generated as shown in Fig. 2 and whose description is given as follows:

- First, rightmost 32 bits of intermediate data is XORed with 32 bits of the subkey generated for that round.
- 32 bits are then transferred to Serpent S-boxes which are basically eight different 4-bit S-boxes that are connected in parallel and process and substitute all the 32 bits of the plaintext.
- Then, bitwise permutation is performed on the 32 bits obtained from the S-boxes.
- Thus, we obtain the round function for a single round. The same process is repeated to obtain the round function for all of the 32 rounds.

## 2.2 ESF Substitution Box

The round function of ESF is based on the SP network. To employ a nonlinear substitution layer in the round function, ESF uses Serpent S-boxes which basically consist of a set of eight different S-boxes {S\_0; S\_1; S\_2; S\_3; S\_4; S\_5; S\_6; S\_7}. Table 1 shows values of these S-boxes in hexadecimal format. These are basically  $4 \times 4$  S-boxes, i.e., they accept input and produce an output of 4 bits. They work on 32-bit intermediate data in parallel with a group of 4 bits for each of the S-boxes. As these are  $4 \times 4$  S-boxes, their cost of implementation is quite less as compared to S-boxes of 8 bits. To process 32-bit intermediate data from the set of  $4 \times 4$  S-boxes,

**Table 1** Serpent S-box used in ESF implementation in hexadecimal format [15]

S0	3	8	F	1	A	6	5	B	E	D	4	2	7	0	9	C
S1	F	C	2	7	9	0	5	A	1	B	E	8	6	D	3	4
S2	8	6	7	9	3	C	A	F	D	1	E	4	0	B	5	2
S3	0	F	B	8	C	9	6	3	D	1	2	4	A	7	5	E
S4	1	F	8	3	C	0	B	6	2	5	4	A	9	E	7	D
S5	F	5	2	B	4	A	9	C	0	3	E	8	D	6	7	1
S6	7	2	C	5	8	4	6	B	D	9	1	F	D	3	A	0
S7	1	B	F	0	D	8	2	B	7	4	C	A	9	3	5	6

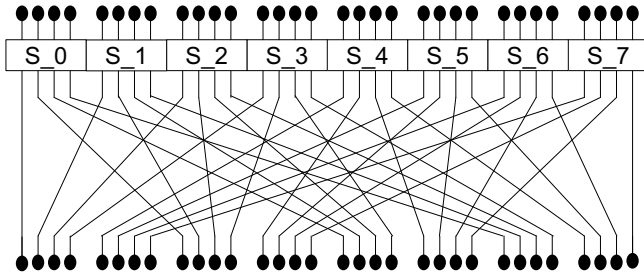


Fig. 3 16-bit Permutation layer for ESF

data is taken in groups of 4 bits and processed in parallel from eight different S-boxes. Thereafter, processed values of S-boxes are sent to the permutation layer.

### 2.3 ESF Permutation Box

32-bit data obtained from the series of Serpent S-boxes are then sent to the permutation layer. Figure 3 shows an ESF P-box (Permutation box). ESF follows the PRESENT permutation layer. It works on bitwise permutation. The following equation shows the working of the PRESENT permutation layer which is followed in ESF as well.

Let “x” be input data, “y” be output data, and “i” denote the byte position. Then the equation to implement bitwise permutation is as follows:

For  $0 \leq i \leq 4$

$$x_{4i} \parallel x_{4i+1} \parallel x_{4i+2} \parallel x_{4i+3} \Rightarrow y_i \parallel y_{i+8} \parallel y_{i+16} \parallel y_{i+24} \tag{1}$$

### 2.4 ESF Key Schedule

ESF accepts key of 80-bit length. For each of the 32 rounds, a subkey of length 32 bits is generated through the key scheduling part of the architecture. For  $i = 1, 2 \dots 31$ , key register K is updated according to the following steps:

- Left Shifting: Values stored in the key register K being shifted left with an offset of 13 bits.
- Bits from position 72–79 are sent to S-box for substitution operation. Serpent S-box0 is used for this purpose:  $[k_{79}k_{78}k_{77}k_{76}] = S0[k_{79}k_{78}k_{77}k_{76}]; [k_{75}k_{74}k_{73}k_{72}] = S0[k_{75}k_{74}k_{73}k_{72}]$
- Key bit from position 43–47 is XORed with a 5-bit round counter.  $[k_{47}k_{46}k_{45}k_{44}k_{43}] = [k_{47}k_{46}k_{45}k_{44}k_{43}] \hat{\ } round\_counter$

- The leftmost 32 bits of the current content stored in the key register “K” are taken as output which is basically subkey “ $K_{i+1}$ ” for that round.

### 3 Hardware Architecture of ESF

Various techniques for the implementation of Serpent S-boxes are proposed. These techniques are also combined with Block Random Access Memory (BRAM) and Read Only Memory (ROM)-based key retrieval methods to evaluate the optimum trade-off between parameters. All these techniques are implemented and evaluated for the round-based design of ESF.

#### 3.1 Using Boolean S-Box

Using Table 1, Sum-of-Product based Boolean expressions of all the eight different S-boxes are obtained using the Karnaugh map technique. With the help of these expressions, S-boxes can be easily implemented using basic gates only, i.e., NOT, OR, and AND. Figure 4 shows the design with Boolean S-boxes implemented in a round-based ESF design. Two 32-bit registers are used for storing the leftmost and rightmost values of the plaintext. An 80-bit key register is used to store the value of subkeys. Each cycle of encryption is completed in a single clock cycle thus giving a

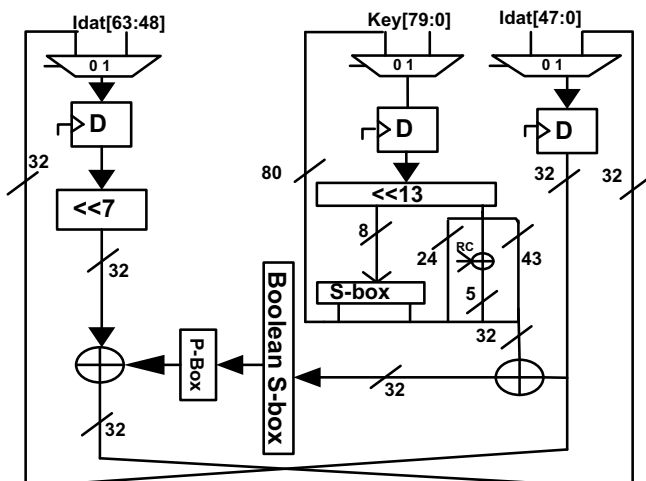
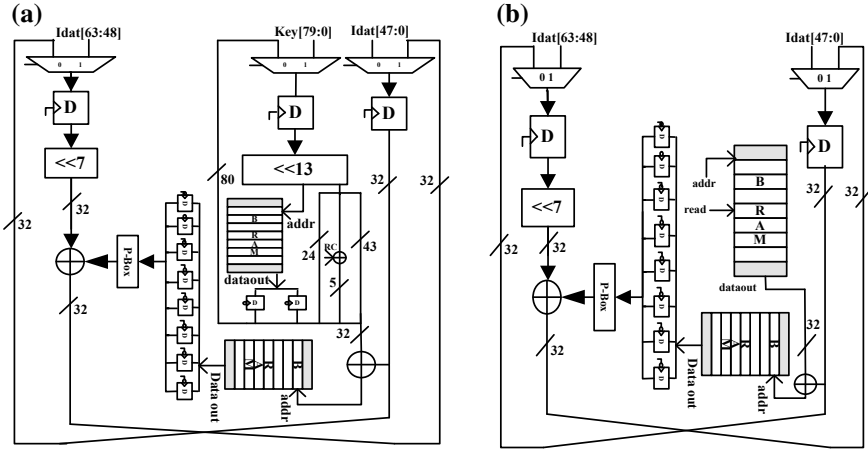


Fig. 4 Round-based architecture of eight-sided fortress





**Fig. 5** Proposed hardware architecture of ESF with **a** BRAM-based S-box **b** BRAM-based S-box and key retrieval mechanism

latency of 32 for 32 rounds of encryption. Since we are using Boolean expressions to design S-boxes, this design is named as Boolean S-box based architecture of ESF cipher.

### 3.2 Using RAM-Based S-Box

The next variation in S-boxes implementation technique is by using BRAM for the implementation of eight Serpent S-boxes. Figure 5a shows the architecture for RAM-based S-box architecture. For all the eight different S-boxes, the contents are stored in BRAM of FPGA instead of using slices because of which, slices available in FPGA are free to be used for any other purpose. Since the implemented S-boxes are RAM-based, extra clock cycles are required to retrieve and process data from the RAM, for which in each round, an extra clock cycle is required, resulting in adding an extra cycle per round, and hence increasing the latency and throughput of the design. BRAM can be used instead of distributed RAM of FPGA for which synchronous S-boxes are designed. In BRAM, each S-box output data is stored as 4-bit hexadecimal data for all 16 values, i.e., each S-box needs  $4 \times 16$  bits of memory. Thus, for 10 S-boxes (two extra S-boxes for key scheduling mechanism), 640 bits of memory is required. Input values of S-boxes are sent to BRAM as address. Data which is stored in that address is basically values which will come after performing the substitution. That value is retrieved and sent to the output port of the RAM. All the values are stored and retrieved in hexadecimal format only. Since S-boxes are also synchronous, extra circuitry is used to aid in controlling and proper functioning of the hardware. To synchronize the design, extra clock cycles are required for each round, thus, increasing the latency of the architecture from 32 to 64.

In the next architecture, BRAM is used for both S-box designing and for retrieval of subkeys for each of the ESF 32 encryption rounds. Figure 5b shows the hardware architecture of the design. For the implementation of this design, only eight S-boxes are required as the key scheduling part is replaced by a RAM module having pre-generated subkeys as stored data. With proper address as input, keys for each of the 32 rounds can be retrieved for the processing. In this way, resource consumption is reduced. But this makes the architecture vulnerable to related key attack [14]. In this architecture, also synchronous S-boxes are used so that FPGA uses only BRAM for its implementation. Thus, extra circuitry is required for controlling the circuit.

The next technique which is studied is by using asynchronous memory module in which BRAM is not used for key storage. Pre-generated subkeys are stored in ROM-based memory module from where they are retrieved for respective rounds while the S-boxes are implemented using BRAMs only. To achieve this, the key schedule is designed for asynchronous operation while S-boxes are designed for synchronous operations. Thus, the architecture uses LUTs and slices instead of BRAM for the implementation of key storage and retrieval part.

## 4 Results

Results are evaluated and compared in two different FPGA platforms, i.e., Spartan-3 and Virtex-4. Comparison of all the architectures is done on the basis of resource consumption which is evaluated by considering the number of flip-flops, LUT, number of BRAM and slices used in the implementation of the architecture in the respective platform. Table 2 shows a comparison of resource consumption for the implementation of the various architectures.

Comparison of the proposed architectures is also done with the PRESENT cipher. In the case of Spartan-3 FPGA, the comparison is done with two architectures of the PRESENT cipher which are designed on basis of different types of techniques used for the implementation of S-boxes namely espresso-based S-box and LUT-based technique for the design of S-box.

Table 3 shows the power requirement for the implementation of various architectures in different FPGA. It also shows the latency and throughput of the architectures for both Virtex-4 and Spartan-3 FPGA. Here also, results are compared with the optimized PRESENT design.

## 5 Evaluation of Results

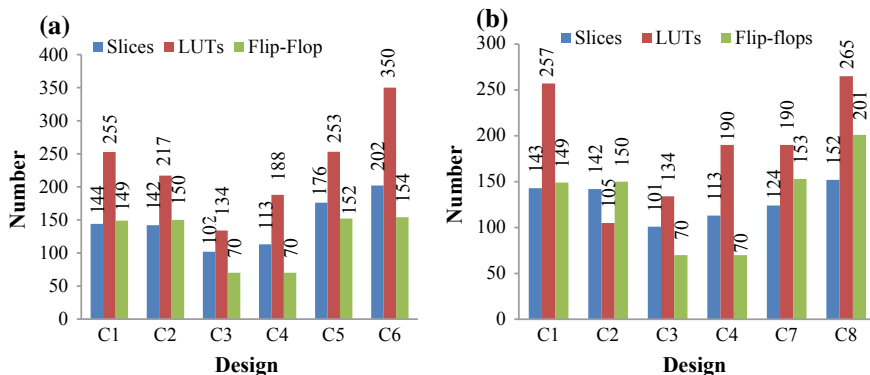
**Resource Consumption:** It is observed that the least number of resources is used when BRAM-based key retrieval mechanism along with BRAM-based S-box (C3) is used. As shown in Fig. 6, the design shows a decrease in the number of LUTs, registers, and slices as intermediate data and S-boxes are stored in BRAM of FPGA.

**Table 2** Comparison of resource consumption for different architectures

Design	State (Bit)	Key (Bit)	BRAM	Slice	LUT	Flip-flop	Fmax (MHz)
<i>xc3s400-5fg456</i>							
Boolean S-box Architecture(C1)	64	80	–	144	255	149	100.040
RAM-Based S-Box Architecture(C2)	64	80	10(64-bits each)	142	217	150	73.292
BRAM-Based Key Architecture(C3)	64	80	9(64-bits each)	<b>102</b>	<b>134</b>	<b>70</b>	166.373
ROM-Based Key Architecture(C4)	64	80	8(64-bits each)	113	188	<b>70</b>	88.833
Espresso-Based PRESENT(C5) [16]	64	80	–	176	253	152	<b>258</b>
LUT-Based PRESENT (C6) [16]	64	80	–	202	350	154	240
<i>xc4vlx25-12ff668</i>							
C1	64	80	–	143	257	149	224.459
C2	64	80	10	142	<b>105</b>	150	152.764
C3	64	80	9	<b>101</b>	134	<b>70</b>	226
C4	64	80	8	113	190	<b>70</b>	172.586
PRESENT-80(C7) [17]	64	80	–	124	190	153	<b>375.66</b>
PRESENT-128(C8) [17]	64	128	–	152	265	201	364.56

**Table 3** Comparison of power and performance in different architectures

Design	State (Bit)	Key (Bit)	Latency (Cycle)	Dynamic power (W)	Static power(W)	Total power (W)	Thr. (Mbps)
<i>xc3s400-5fg456</i>							
C1	64	80	32	Negligible	0.060	0.060	200.08
C2	64	80	64				73.291
C3	64	80	64				166.373
C4	64	80	64				88.833
C5	64	80	32				<b>516</b>
C6	64	80	32				480
<i>xc4vlx25-12ff668</i>							
C1	64	80	32	0.020	0.233	0.253	<b>448.918</b>
C2	64	80	64	0.022	0.331	<b>0.353</b>	152.764
C3	64	80	64	0.012	0.331	0.343	226
C4	64	80	64	0.012	0.331	0.343	172.586
C7	64	80	32	0.012	0.232	0.244	180.77
C8	64	128	32	0.015	0.233	0.248	171.56



**Fig. 6** Results obtained for different architectures in **a** Spartan-3 FPGA **b** Virtex-4 FPGA

Also, this technique gives the best maximum operational frequency among all architectures for both the FPGA. As compared to LUT-based PRESENT (C6), C3 shows 49%, 62%, and 56% reduction in slices, LUTs and registers, respectively. For both FPGA, C3 requires the least resources followed by C4, C2, and C1 as data and S-boxes are stored in BRAM. In FPGA, ROM is designed using LUTs and slices, so on changing the key schedule from BRAM-based (C3) to ROM-based (C4), the number of LUTs and slices are increased, while registers remain the same. It should be kept in mind that C3 and C4 are prone to side-channel attack.

**Performance and Power Consumption:** Implementation in Spartan-3 FPGA requires less power than Virtex-4. For Spartan-3, all designs require the same power. In Virtex-4, least power is required by C1, followed by C3 and C2. C3 and C4 require the same power consumption. To synchronize the circuit, the latency of the architecture is increased in case of BRAM-based designs (C2, C3). For Boolean S-box design (C1), latency is low and resource utilization is more. Throughput (thr.) depends on latency and maximum operational frequency (FMax.) of the design. It also varies with the FPGA device used for the implementation. Best throughput is given by C1 when implemented in Virtex-4 FPGA, followed by C3, C4, and C2. While for Spartan-3 FPGA, design C5 and C6 give good results followed by C1, C3, C4, and C2.

## 6 Conclusion

ESF is a lightweight Feistel cipher with SPN round function. In this work, different S-box implementation techniques are presented and evaluated. The designs are implemented in different FPGAs and results are compared on the basis of resource consumption and power requirement. Best results are obtained when BRAM is used

for the implementation of both S-boxes and key retrieval mechanism (C3). For Virtex-4 FPGA, C3 required the least resources and had good throughput results. C1 gives the best throughput in Virtex-4 platform. This results in reduction of LUTs, slices, and registers. These resources can be used for some other purpose in the same FPGA device.

Further studies are required to study the relationship between latency and area of the architecture. There is room for designing of other architectures of ESF like pipelined, serial, parallel, etc. This will lead to the establishment of the trade-off between different resource requirements and performance. A detailed evaluation will be required to study resourcefulness of the design.

## References

1. Xiang, Z., Zhang, W., Bao, Z., & Lin, D. (2015). RECTANGLE: a bit-slice lightweight block cipher suitable for multiple platforms. *Science China Information Sciences*, 58, 1–15.
2. Standard, D. E. (1977). Federal information processing standards publication 46. *National Bureau of Standards, US Department of Commerce*, vol. 23.
3. Pub, N. F. (2001). 197: Advanced encryption standard (AES). *Federal Information Processing Standards Publication, 197*, 0311.
4. Kong, J. H., Ang, L. M., & Seng, K. P. (2015). A comprehensive survey of modern symmetric cryptographic solutions for resource-constrained environments. *Journal of Network and Computer Applications*, 49, 15–50.
5. Bogdanov, L. R. Knudsen, G. Leander, C. Paar, A. Poschmann, M. J. Robshaw, Y. Seurin, & Vikkelsoe, C. (2007). PRESENT: An ultra-lightweight block cipher. In *International Workshop on Cryptographic Hardware and Embedded Systems*, vol. 4727, pp. 450–466.
6. Knudsen, L., Leander, G., Poschmann, A., & Robshaw, M. J. (2010). PRINTcipher: A block cipher for IC-printing. In *International Workshop on Cryptographic Hardware and Embedded Systems*, pp. 16–32.
7. Hong, D., Sung, J., Hong, S., Lim, J., Lee, S., Koo, B. S., Lee, C., Chang, D., Lee, J., Jeong, K., & Kim, H. (2006). HIGHT: A new block cipher suitable for the low-resource device. In *International Workshop on Cryptographic Hardware and Embedded Systems*, pp. 46–59.
8. Guo, J., Peyrin, T., Poschmann, A., & Robshaw, M. (2011). The LED block cipher, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Preface, vol. 6917, pp. 326–341.
9. Wu, W., & Zhang, L. (2011). LBlock: a lightweight block cipher. In *International Conference on Applied Cryptography and Network Security*, pp. 327–344.
10. Gong, Z., Nikova, S., & Law, Y. W. (2011). KLEIN: A new family of lightweight block ciphers. In *International Workshop on Radio Frequency Identification: Security and Privacy Issues*, pp. 1–18.
11. Xuan, L. I. U., Zhang, W. Y., LIU, X. Z., & Feng, L. I. U. (2014). Eight-sided fortress: a lightweight block cipher. *The Journal of China Universities of Posts and Telecommunications*, 21, 104–1282014.
12. Hell, M., Johansson, T., & Meier, W. (2007). Grain: A stream cipher for constrained environments. *International Journal of Wireless and Mobile Computing*, 2, 86–93.
13. De Canniere, C. (2006). Trivium: A stream cipher construction inspired by block cipher design principles. In *International Conference on Information Security*, pp. 171–186.
14. Li, L., Liu, B., & Wang, H. (2016). QTL: A new ultra-lightweight block cipher. *Microprocessors & Microsystems*, 45(PA), 45–55.

15. Biham, E., Anderson, R., & Knudsen, L. (1998). Serpent: A new block cipher proposal. In *International workshop on fast software encryption*, pp. 222–238.
16. Sbeiti, M., Silbermann, M., Poschmann, A., & Paar, C. (2009). Design space exploration of present implementation for FPGAs. In *Programmable Logic, 2009. SPL. 5th Southern Conference, IEEE, Sao Carlos*, pp. 141–145.
17. Lara-Nino, C. A., Diaz-Perez, A., & Morales-Sandoval, M. (2017). Lightweight hardware architectures for the PRESENT cipher in FPGA. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 64, 2544–2555.

# Two-Dimensional Hybrid Authentication for ATM Transactions



M. F. Mridha, Jahir Ibna Rafiq and Wahid Uz Zaman

**Abstract** Advancement of information technology leads toward a world with process automation to perform a task more efficiently and avail the services with ease. Banking sectors are not an exception and are moving from traditional manual banking system to an electronic entity. The basic functionality of a bank, out of many, is to deposit money into user accounts and retrieve as per account holder's necessity. However, as time is precious, eventually account holders may not expect to spend too much time in the queue for depositing or retrieving their money. That is why the need for ATM comes into the picture to make the user's life easier. However, it comes with some questionable possibilities for false attacks as well. Thus, a proper user authentication mechanism is needed to overcome these fraudulent activities. Our proposed method gives a new dimension to this authentication which is a hybrid version of an existing authentication system for the ATM transaction by using a Graphical pattern password along with current PIN code supplied from the bank. This Graphical password is a version, which has been invented by Google's Android pattern unlock system. In our proposed mechanism, we combine both Graphical pattern and PIN and incorporated security to enhance reliable transactions. More specifically, the secret encryption key is generated from a PIN using the PRESENT algorithm. Finally, the ciphertext is created using digit stream from the Graphical pattern and secret encryption key. This hybrid process to detect intrusion will significantly enhance security. Our primary focus is to develop a robust and flexible user authentication system to avoid common authentication problems. The proposed approach needs no additional hardware and device dependency.

**Keywords** ATM transaction · User authentication · Graphical password · PIN code · Encryption · ATM security

---

M. F. Mridha · J. I. Rafiq (✉) · W. U. Zaman  
University of Asia Pacific, 74/a, Green Rd, Dhaka 1215, India  
e-mail: [Jahir@uap-bd.edu](mailto:Jahir@uap-bd.edu)

M. F. Mridha  
e-mail: [firoz@uap-bd.edu](mailto:firoz@uap-bd.edu)

W. U. Zaman  
e-mail: [Wahid@uap-bd.edu](mailto:Wahid@uap-bd.edu)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_19](https://doi.org/10.1007/978-981-15-0694-9_19)

## 1 Introduction

In the last two decades, information technology has grown in a very vast scale, as an effect of using computerization on a mass level. The technological innovation along with the optimized cost drives us toward the massive growth of the computer system, and thus digitalization is involved in every sector. As time passes by, computer processing power is increasing, and subsequently, prices are reduced along with physical sizing, resulting in computer processing available on every possible field from medical science to space discovery. This digitalization improves our way of living and helps us to compute more things within the shortest possible time. As a result, the manual working system replaces the computerized machine-enabled systems to serve more than the traditional systems. However, similar to other good systems, there are some dark sides of the digital system which also become available to the people who are using this digital system to fulfill their evil wishes. In the real world to protect against crime, law-enforcement agencies work relentlessly. Similarly, contingency plans are taken in case of a compromise of data in the virtual world.

The financial sector is one of the critical areas where the effect of digitalization can observe visibly. For example, in the traditional banking process, a massive number of accounting books needs to keep track of the transactions, every time an accounting book needs to open manually to update transaction information for a particular account. It was a severe time-consuming process; imagine a branch with 1000 account holders; each requires at least 10 of this kind of accounting books to keep all accounts related information. So, whenever someone is depositing/withdrawing from the account, it requires to find that particular accounting book and specific accounts to update a row. The whole process of finding just one account usually involves multiple people to work, and they have to synchronize in between to serve people efficiently. Also, human error factors in updating information were a big concern along with isolation of that error. So, especially account searching and updating transaction information in the new digital system gives comfort to the administrators and their respective users. Now a bank officer with a computer can answer more than 100 customers in a day not only about the account balance but also about all the information available on that account. Central banking server and banking system automation together achieve this efficiency. Another useful feature of the banking digitization is the invention of Automatic Teller Machines (ATM) which makes a banking transaction live on a  $24 \times 7$  basis and facilitates their account holders in such a way that in case of emergency, irrespective of day or night time, an account holder can withdraw money as per their predefined account limit. With the advancement of ATMs, now cash deposit into those ATMs are also available.

Despite all these excellent features, there are some dark sides of the automatic banking system which are being used by some fraudsters to steal money from accounts. These fraudsters attack the digital system in many ways and try to fool the system so that the system can consider those imposters as legitimate account



holders for the transaction. ATM fraudulent transactions [1] are increasing day by day because from ATM we can receive hard cash which cannot be tracked later.

Several remediation steps have been taken [2] to improve user authentication. Our proposed method uses a hybrid approach to authenticate proper banking users. This method includes an existing PIN code along with a newly developed Graphical pattern system which is inspired by Google's Android pattern unlock system [3]. Like Google's Android pattern, we have taken a large  $4 \times 4$  grid [4, 5] plotted with dots on the screen. To take user input, we have considered a touch-based input panel. Our proposed system tweaks the backend calculation by assigning a fixed number into each dot and resolves the user input pattern into a stream of digits. Users then input their existing PIN code supplied by their respective bank on the input panel which is considered as an initial key for asymmetric encryption [6] algorithm, PRESENT, to generate a secret key and apply encryption process to that digit stream before transmitting it to the bank server. Bank server will then again reconstruct the digit stream from the coded received message by using the same encryption algorithm and prestored PIN code from the particular account. This derived digit stream is then checked against the stored digit stream on that account and allows/denies further transaction upon the results.

The rest of the paper is organized as follows. Section 2 provides the related research works on a particular field of interest (i.e., ATM user authentication and security). Section 3 describes the proposed method briefly: A overview of the whole method, B Steps required to perform the whole authentication process, C Graphical pattern system, D Encryption algorithm, and E Results and benefits of the proposed method. Finally, Sect. 4 concludes the paper with future works.

## 2 Related Work

Many research works have been carried out due to growth and acceptability of E-banking system toward its users to make transaction secure. In our study, we have seen that many different kinds of account holder authentication systems have been proposed to enhance the identification of legitimate account holders. These authentication methods are mainly based on Biometrics like Fingerprint, Iris, Face detection, and Vein [7–10]. Despite the effectiveness of these biometric authentication techniques, there is also a chance of copying these parameters by an imposter. So, later researchers found that a single authentication system is not good enough to achieve the goal and thus multifactor authentication has been introduced [2, 11]. These multiple authentication processes force users to enter two or more authentication parameters during a transaction to complete the account holder validation. Some of the authentication parameters are prestored during the account opening stages like PIN code, Biometrics (Fingerprint, Face, Iris) and some parameters are produced and supplied in real time during the transaction (like one-time code through the token devices, smart cards, and SMS on the phone).

Also, the increasing number of IoT devices can lead toward a new possibility of a user's authentication system, sometimes even without using the ATM cards [12]. There are cell phones, smartwatches, smart goggles, or even car keys that can be used as a part of token-based user authentication. So, we assume that in the near future, the old traditional PIN code based authentication or single authentication system is going to be extinct and new fusion systems with a combination of Biometrics, IoT devices, and other forms of authentication parameter will be added for authentications. As a consequence, the main focus of this paper is to develop a robust and flexible user authentication system to avoid common authentication problems. The proposed approach needs no additional hardware to carry and requires no device dependency.

### 3 Proposed Methodology

#### 3.1 Short Summary of Proposed Architecture

In this paper, we propose a hybrid approach for the user authentication method unlike the single PIN code based inputs or Biometrics input. We proposed a user authentication system which is a combination of two different user inputs. For doing this, we have considered the ATM with a touch screen based user input panel. Although we know that keypad-based ATMs are the most common systems in use, due to the technology advancement, availability, and low price of touch screen based devices, they are growing in numbers and thus ATMs with a touch panel are also increasing day by day. Our authentication method combines two-level user inputs which have two parts. First, users have to input the Graphical pattern password which should be similar to drawing a pattern by connecting some dots on ATM screen (similar to screen password on Android mobile phones). For collecting user inputs, we are considering a  $4 \times 4$  grid and mapping it with a predefined number for each cell.

#### 3.2 Details of Functional Steps

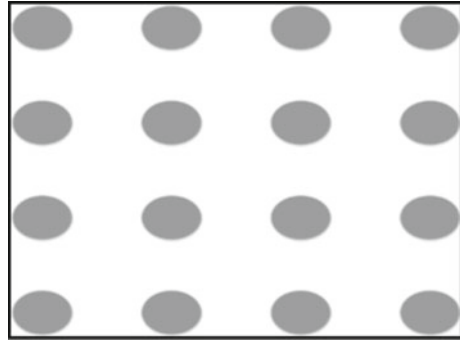
The following steps are derived from Fig. 1 to describe the whole authentication process.

**Step-1:** User swipes their Debit/Credit card into the ATM machine card slots and the ATM reads and collects necessary user details from the card.

**Step-2:** Information gathered from the card will be sent back to the card provider merchant (e.g., Visa/MasterCard/Amex) for identifying the actual recipient bank server.

**Step-3:** Recipient bank server will check the card information and account details for accepting further inputs and thus notifying the particular ATM machine about

**Fig. 1** Graphical pattern password initial input panel



further user inputs. For stolen card/expiry card/account disabled case, the bank server will send a denied message to show proper notification for the users.

**Step-4:** Upon getting the initial account validity from the bank server, the ATM machine will prompt for entering the first phase Graphical pattern password, which will show a  $4 \times 4$  grid with dots connected by touching one dot to another to create a pattern.

**Step-5:** Grid conversion function will translate the pattern into a digit stream where each of the grid cells will be mapped by a predefined two-digit decimal number. The digit stream is then stored locally on the ATM machine on a temporary basis and moves forward for the next step inputs.

**Step-6:** In this phase, second dimensional user input is taken which is nothing but the old PIN code (4/6 digit) assigned by the bank server. This PIN code is also a 4/6 digit decimal number.

**Step-7:** Now an encryption key generation function will generate the key. In our case, we have considered the PRESENT symmetric cryptography algorithm along with 80 bits key size.

This same algorithm is also used later on to encrypt the message with the secret key generated from the PIN code inputs from the users before sending it to the respective bank server.

**Step-8:** The PRESENT encryption algorithm will encrypt the previously stored digit stream which is converted from Graphical pattern inputs by using the secret key derived from the previous step. The PRESENT algorithm will consider a 64-bit block at a time

to apply encryption and will start sending the coded message through the ATM network infrastructure toward the bank server.

**Step-9:** The bank server will receive the transmitted packet from the ATM and get the fragmented encrypted bits. After collecting all the fragmented payload, the bank server will start decrypting the payload using the same secret key as generated by the same encryption PRESENT algorithm through the use of stored PIN code from that particular user account.

**Step-10:** The bank server will reconstruct the digit stream from the encrypted message sent earlier by the ATM machine and check with the prestored digit stream

on that particular account. Here, we have considered the stored digit stream which was entered during the initial account opening process by the user himself. During the account opening time, the user will input their own Graphical pattern password through any Tablet/Touchable user interfacing device from designated branches and that Graphical pattern password is then converted to a digit stream for storing on the respective user account for future authentication.

**Step-11:** The bank server will validate the user authentication by sending allow notification to the ATM machine for further transaction inputs or sending denied message to terminate the particular session if those two-digit streams do not match with each other.

In this architecture, the existing network infrastructure for ATMs will be used and there is no need for any change on the backend connections. We have considered that during the account opening, the respective bank branches will collect the user Graphical pattern password by using any touchable input devices and also to avoid common guess on Graphical password constraints like a minimum of five touch long-pattern password will allow or include minimum one overlapping on cell, etc. For the ATM machine, we have considered the touch-based screen as a user input panel with traditional card swiping slots, cash dispensers, print paper outlets, etc.

### ***3.3 Graphical Pattern User Password System***

In our proposed authentication system, we have combined two user inputs: Graphical pattern drawing with the existing PIN code. These Graphical pattern password input systems are derived from the invention of the Graphical password-based authentication from Google's Android Pattern un(lock) system [3]. The concept was to draw something on the screen which can relate to some fixed coordinates of pixels acting as a point of touch (looks like dots symbols to represent) to create a pattern which will be matched against the stored one.

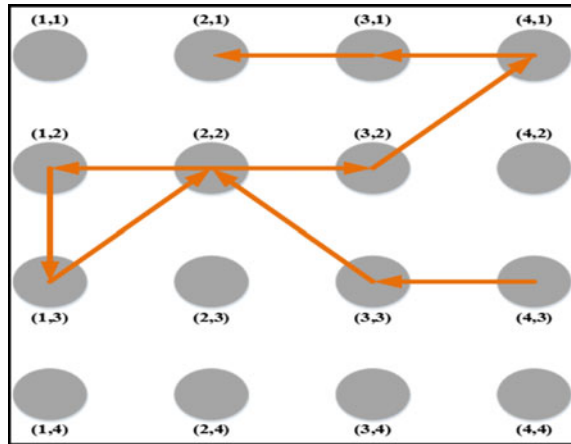
Here, in this paper, we are using a similar Graphical system but we have tuned the underlying mechanism in a different way which is flexible and reprogrammable for changing.

In Fig. 2, a grid with dots on specific pixels has been depicted. Here, each dot contains some fixed pixels from the screen, plotted with round symbols like a dot and is represented with two-dimensional coordinates (x, y). Between one dot to another, there are some fixed gaps which will help to draw a line by connecting dots in an order to create a pattern.

### ***3.4 Lightweight on-Premise Data Encryption System***

To maintain the confidentiality of the pattern password, we are proposing encryption method upon the digit stream which has been derived right after the pattern drawing.

**Fig. 2** Drawing a pattern password on the input screen



For data encryption, we have considered a symmetric encryption method which is lightweight because we are only dealing with a series of numbers and considering processing on-premises ATMs which have limited processing capacities. To achieve this, we considered PRESENT as the encryption algorithm [13], which is a kind of block cipher category. The PRESENT algorithm developed by Orange Labs in 2007 is 2.5 times smaller than the AES algorithm [13].

Symmetric encryption algorithm requires the same key for both encryption and decryption process. However, there is a significant concern about how the encryption key is delivered through the network as it is not safe if someone can steal the key during transmission. That is why we considered a way to prevent this. Using the existing PIN code supplied from the bank to the specific user acts as a source of an initial key which is reused with the PRESENT encryption algorithm generates 80-bit secret key. With this newly derived secret key, the digit stream which was derived earlier can be encrypted. As the same PIN code is prestored with the particular user account, the bank server can regenerate the secret key and decrypt the coded message transmitted from the ATM with that particular user account. So, here the account PIN code is considered as a session key for a particular transaction to regenerate the real encryption secret key. The PRESENT algorithm works with a 64-bit block size with a round function iterated 31 times. Each round consists of three sub-functions: addRoundKey, pLayer, sBoxLayer. addRoundKey is XOR of 64-bit round key and the state. A pLayer is a bitwise permutation. sBoxLayer is 64-bit nonlinear transformation.

### 3.5 Determine the Strength of a Pattern

Distance metrics should be clearly outlined. For our work, we have assumed one horizontal or one vertical movement is weighted as one. Similarly, two such steps

are counted as two. For, diagonal distance, we merely apply Pythagorean theorem. Hence, one diagonal distance is the square root of  $2 = 1.41$ , and two diagonal distances are the square root of  $8 = 2.83$ . For a Knight in Chessboard, we have one diagonal and one horizontal or one vertical step such that the weight is  $(1.41+1)$ . Let us consider a pattern of nine digits ( $5 \rightarrow 1 \rightarrow 9 \rightarrow 4 \rightarrow 6 \rightarrow 7 \rightarrow 3 \rightarrow 8 \rightarrow 2$ ). Using this weighting technique, we get 17.7. So the higher the weight we get, the better we secure our pattern.

Our security can be breached through shoulder surfing, phishing, or guessing. Because the first falls within the user's responsibility, we consider guessing a threat. We have used the Divide and Conquer Algorithm to test the strength of our pattern, which is a recursive function to determine the correct number. However, the more guesses it takes, the stronger our pattern is. The recursive function is summed up in the following diagram. Because the pattern is a variable and to compromise the whole pattern, intruders have to decode each digit in one step; we can quickly enhance the robustness of our pattern.

In our proposed method, some underlying benefits can be marked which help the bank to identify its rightful account holder. Our main focus was to develop a flexible format of user authentication which can be quickly adopted and would require minimal changes from the existing system to operate. Below are some salient points which show the proposed method's benefits regarding the flexibility and usefulness.

1. Easily reprogrammable to change the mapping function, which can give privilege to the banks to redesign as per their demand and makes the mapping different/unique from one bank to another. So, just changing the number allocation makes the resulting digit stream completely different.

2. Digit stream generation is so flexible that a user can choose any size of the pattern which in turn only requires a few digits of numbers to store. For example, a pattern with 50 consecutive touches produces only 100-digit long number stream to store.

3. In research, it has been shown that with a  $4 \times 4$  Grid, we can have 4,350,069,823,024 number of the possible patterns [5]. So, a wide variety of patterns is possible and not just 4 or 6 digits long PIN code to choose.

4. During account opening, a user enters the pattern which is just like their password. So, it provides better comfort to the user regarding authenticity instead of only PIN code which is generated and supplied by bank servers.

5. A crucial secret generation with the existing PIN code will not only reduce the risk of key transmission but also does not need any changes at the existing banking system.

6. Using an on-premises encryption method can maintain the integrity of user data which is helpful if the ATM network does not facilitate by default encryption on data transmission.

## 4 Conclusion

Banking transaction system becomes digital by the use of ATM instead of disbursing money from the bank directly. This digitalization of manual money withdrawal process possesses a severe security risk of fraudulent transaction and possibilities of non-authentic users' access. As the system is automated, it is not possible to identify the individuals by asking their identity or by observing the activities/body languages of a non-authentic user. Proper identification is achievable through human intervention. So the effort should be focused to make the authentication challenges multiples and confidential/ challenging to guess or replicate. In this regard, we propose to use a different mechanism for authentication which aggregates the existing authentication parameters with the new system to create a hybrid system. We firmly believe that our proposed mechanism enhances the banking transaction to be more reliable and secure.

This system may also be implemented with the Internet banking system/mobile banking system and might be integrated with another new kind of authentication systems like NFC-based devices or IoT-based devices which can remove the need of using a plastic ATM card. Also, the proposed method should explore with the POS transactions system. Some of the shared Graphical password risks like shoulder surfing, natural guess ability should be carefully considered to make the method more robust and efficient before deploying at a mass level. Finally, some critical issues regarding the performance, user adaptability and usability, and security of the system need to be explored.

## References

1. Bank Fraud & ATM Security website [Online]. <http://resources.infosecinstitute.com/bank-fraud-atm-security/#gref>.
2. Muhammad-Bello, B. L., Alhassan, M. E., & Ganiyu, S. O. (2015). An enhanced ATM security system using second-level authentication. *International Journal of Computer Applications*, *111*(5), 8–14.
3. Jermyn, I., Mayer, A., Monrose, F., Reiter, M. K., & Rubin, A. D. (1999). The design and analysis of graphical passwords. In *Proceedings of the 8th USENIX Security Symposium*, pp 2–12.
4. von Zezschwitz, E., De Luca, A., & Janssen, P. (2015). Heinrich Hussmann: Easy to draw, but hard to trace? On the observability of grid-based (Un)lock patterns. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 2339–2342.
5. Aviv, A. J., Budzitowski, D., & Kuber, R. (2015). Is Bigger Better? Comparing user-generated passwords on  $3 \times 3$  vs.  $4 \times 4$  grid sizes for android's pattern unlock. In *Proceedings of the 31st Annual Computer Security Applications Conference*, pp. 301–310.
6. An Overview of Cryptography website. [Online]. Available: <http://www.garykessler.net/library/crypto.html>.
7. Iyabode, A. M., Nureni, Y. N., Adebayo, A. F., & Olamide, O. A. (2015). Card-less electronic automated teller machine (EATM) with biometric authentication. *International Journal of Engineering Trends and Technology*, *30*(1), 99–105.

8. Vats, H., Ruhl, R., & Aghili, S. (2015). Fingerprint security for protecting EMV payment cards. In *The 10th International Conference for Internet Technology and Secured Transactions*, pp. 95–100.
9. Sui, Y., Zou, X., Du, E. Y., & Li, F. (2014). Design and analysis of a highly user-friendly, secure, privacy-preserving, and revocable authentication method. *IEEE Transactions on Computers*, 63, 902–906.
10. Dileepsai, Y., & Sudarvizhi, S. (2016). Card less access to POS transactions. *International Journal of Applied Engineering Research*, 11(7), 5231–5236.
11. Awotunde, J. B., Jimoh, R. G., & Matiluko, O. (2015). *Emmanuel: Secure Automated Teller Machine (ATM) Using Fingerprint Authentication and Short-code Message in a Cashless Society*. Published in Academia.edu, pp 1–12.
12. Abdulrahman Alh othaily, Arwa Alraw ais, Xiuzh en Cheng, Rongf ang Bie: A Novel verification method for payment card systems. *Personal and Ubiquitous Computing*, 19(7), 1145–1156.
13. Bogdanov, A., Knudsen, L. R., Leander, G., Paar, C., Poschmann, A., Robshaw, M. J. B., Seurin, Y., & Vikkelsoe, C. (2007). PRESENT: An ultra-lightweight block cipher. In *CHES: International Workshop on Cryptographic Hardware and Embedded Systems*, Vol. 4727, pp. 450–466.



# **Intelligent Computing Techniques**

# Artificial Neural Network Based Load Balancing in Cloud Environment



Sarita Negi, Neelam Panwar, Kunwar Singh Vaisla  
and Man Mohan Singh Rauthan

**Abstract** With heavy demand for cloud technology, it is important to balance the cloud load to deliver seamless Quality of Services to the different cloud users. To address such issues, a new hybridized technique Artificial Neural Network based Load Balancing (ANN-LB) is introduced to calculate an optimized Virtual Machine (VM) load in cloud systems. The Particle Swarm Optimization (PSO) technique is used to perform task scheduling. The performance of the proposed ANN-LB approach has been analyzed with the existing CM-eFCFS, Round Robin, MaxMin, and MinMin algorithms based on MakeSpan, Average Resource Utilization, and Transmission Time. Calculated values and plotted graphs illustrate that the presented work is efficient and effective for load balancing. Hybridization of ANN and iK-mean methods obtains a proper load balancing among VMs and results have been remarkable.

**Keywords** Cloud computing · ANN · iK-mean · Clustering · Load balancing

## 1 Introduction

In the research computing world, (in 2000) a new trend Computing Technology has arrived to change the working approach of computer and Internet users. From the history of utility computing (or Computer utility, a computing resource package that includes computation, services, storage, and computer resource in rent), researchers

---

S. Negi (✉)

Uttarakhand Technical University, Dehradun 248007, Uttarakhand, India

e-mail: [sarita.negi158@gmail.com](mailto:sarita.negi158@gmail.com)

K. S. Vaisla

B.T.KIT, Dwarhat, Uttarakhand, India

S. Negi · N. Panwar · M. M. S. Rauthan

SOET, HNBGU, Srinagar, Garhwal 249161, Uttarakhand, India

have found that cloud is the most sophisticated, reliable, and service-oriented technology. *Abstraction* and *Virtualization* are the two major concepts of the cloud. In *Abstraction*, the cloud hides implementation information from the user as data stored in locations that are unknown and in *Virtualization*, resources are pooled and shared by various users on the metered basis [1].

Uncontrolled growth of DataCenters (DC) may lead to a lack of availability of resources. In such case, a load balancing policy can handle the cloud resources and make resources available to each Host and VM. Scheduling of tasks and load balancing in virtual machines (VMs) is the NP-hard problem hence a suitable scheduling technique is required to solve VM allocation problems [2]. The data that cloud provides to the user must be scheduled and balanced among VM and Hosts [3]. Live migration of VM is a better approach to slow down the processing cost [4] whereas cloud computational cost can be minimized through proper utilization of resources [5]. Utilization of resources may be achieved through proper mapping and allocation of tasks to VM and VM to Host [6]. Scheduling of tasks to single user and multi-user has different aspects of delay bounds for the tasks. Such delay bounds can be deduced by using offloading policy in edge cloud [7]. VM and task scheduling in cloud computing provide flexibility, scalability, load sharing, etc. Proper scheduling of resources improves load balancing such as VM migration [8, 9] and task migration [10, 11]. VM and task migration can have great impact on cloud performance. The various approaches of task migration techniques are non-live migration, post-copy, live migration, pre-copy, and triple TPM etc., [10]. VM live migration and its impact should be low as higher migration increases the processing cost [12].

Management of cloud resources demands the implementation of efficient load balancing techniques. Various categorizations of load balancing have been defined for the traditional computing environment: Static, Dynamic, and Mixed Load Balance [8]. As cloud provides the high mobility of nodes, the spatial distributed nodes need to be balanced using Centralized, Distributed, and Hierarchical Load Balancing methods [13]. This paper focuses on the concept of soft computing based technique, viz, Artificial Neural Network (ANN). The objective of the research is to introduce the process initiated by the VM manager using ANN-based Back Propagation Network (BPN) method. BPN calculates the load of each available VMs and improved K-mean (iK-means) clustering method performs clustering of VMs into underloaded VMs and overloaded VMs. Incoming user tasks that are admitted to cloud at runtime are allocated to underloaded VMs using the PSO algorithm.

The organization of the paper is as follows: Previous work on load balancing in the cloud environment has been discussed in Sect. 2. Section 3 highlights the proposed system model. Section 4 elaborates the implementation of the proposed model while Sect. 5 explains the evaluated outcomes of the proposed work. Finally, the conclusion and future scope are covered in Sect. 6.

## 2 Literature Review

Many researchers worked in load balancing to improve Quality of Service (QoS) of cloud computing. The researches have suggested different methods for load balancing. In this section, various load balancing algorithms are discussed in detail.

Hamsinezhad et al. [10] add up task and VM migration schemes to achieve efficient load balancing in a cloud environment. The work has shown the migration methods on task by combining Yu-Router and Post-Copy migration methods. The algorithm decreases the migration time, overhead, and transmitted data rate. The authors have explained migration in a mesh network that partitions the network into the subnetworks ( $P_{sub}$ ). The migration of  $P_{sub}$  is given in Eq. (1). Number of stages ( $S$ ) for the migration of subtasks to the D.M is calculated using Eq. (2).

$$p_{sub} = (d/p \times w/q \times h/r) \quad (1)$$

$$S = \max(d/p \times w/q \times h/r) \quad (2)$$

where the size of the network is represented by ( $p \times q \times r$ ) and ( $d \times w \times h$ ) is the number of nodes distributed on the network. The research work of [10] reduces task transmission time and data overhead but delay overhead is still a drawback of the algorithm. These drawbacks motivate to introduce a new approach that enhances the transmission time of tasks.

To minimize the workload between servers, an application live migration method has been introduced for large-scale cloud networks [14]. The application live migration takes place by three events, i.e., workload arrival, workload departure, and workload resizing (varying resource size). Li et al. [14] introduced a concept of workload in an encapsulation of application and the underlying operating system of VM. The server node that is running VM is referred to as open box and the server node lacking of VMs is referred to as close box. The arrival of workload is further assigned to an open box. The work shows “*how application (task) migration can perform remapping of workloads to the resource node*”. This migration reduces the number of open boxes. The size of workload has been divided into subintervals 2 M-2 and is represented in levels. The approach seems to be energy efficient but large number of migrations can lead to high processing time. The use of three different algorithms, i.e., workload arrival, workload departure, and workload recycling may increase complexity of the network. The proposed approach reduces complexity by introducing supervised learning approaches for load balancing.

Devi et al. [15] introduced an Improved Weighted Round Robin (IWRR) load balancer where all the tasks are assigned to the VMs according to the IWRR scheduler. After completion of each task, the IWRR load balancer checks if there is a need for load balancing. If the number of tasks assigned to VM is higher, then the IWRR load balancer identifies VMs load. IWRR estimates the possible completion time of all

tasks assigned to that VM. The number of task migrations is significantly reduced in IWRR load balancer due to widespread identifying of the most suitable VM for each task. When overloaded VM drops below its threshold value, the task can be migrated from overloaded to underloaded VM. In order to identify the VMs having the highest and lowest load, load imbalance factor is calculated using the sum of loads of all VM, load per unit capacity (LPC), and threshold ( $T_i$ ), which are defined in Eqs. (3), (4), and (5), respectively.

$$L = \sum_{i=1}^k l_i \quad (3)$$

$$LPC = \frac{L}{\sum_{i=1}^m c_i} \quad (4)$$

$$T_i = LPC \times C_i \quad (5)$$

where the number of VMs in a DataCenter (DC) is represented by  $i$  and  $C_i$  represents the node capacity. VM load imbalance factor is defined by

$$VM \text{ load} \begin{cases} < T_i - \sum_{v=1}^k l_i, & \text{Underloaded} \\ > T_i - \sum_{v=1}^k l_i, & \text{Overloaded} \\ = T_i - \sum_{v=1}^k l_i, & \text{Balanced} \end{cases} \quad (6)$$

The drawbacks of the reviewed literatures motivate to introduce a new method of finding VM load using intelligence artificial neural network method. The obtained load is further clustered into underloaded and overloaded VMs; thus the tasks are assigned to underloaded VMs. The introduced model focuses to enhance resource utilization, transmission time, and makespan.

### 3 System Model and Proposed Work

#### 3.1 Artificial Neural Network Based Load Balancing (ANN-LB)

In this system model, it is assumed that there is a set of a physical machines  $PM = (PM_1, PM_2, \dots, PM_M)$  where each PM holds the set of virtual machines  $VM = (VM_1, VM_2, \dots, VM_j)$ . For the execution, a number of tasks ( $t_1, t_2, \dots, t_i$ ) are assigned to VMs, respectively. VMs use their resources and run parallelly and independently. Load balancing has always been necessary to remove imbalance execution of a task.

The heavy load on the current VMs leads to unbalanced DCs and resource underutilization. Such issues can be resolved by introducing the clustering process on VMs in each PM based on current load of VMs. Based on the load, VMs are grouped.

### 3.1.1 Back Propagation Network (BPN)

The Back Propagation Network (BPN) is used to support several VMs all together for load calculation with the aim to reduce clustering time. A supervised learning Artificial Neural Network based BPN is one of the best neural approaches. Rumelhart, Hinton, and Williams introduced the BPN in 1986. It has the facility to propagate errors toward the back from the output layer units to the hidden layer units.

The role of BPN is to calculate the load of VMs which is further realized by improved K-means (i-Kmean) for VM clustering. In the first step of the algorithm, all VMs with their information are fed into the BPN to evaluate their current processing load on VMs. BPN algorithm performs weight calculation during the learning period of the network. It works on different phases: input  $A_i$  feed-forward, error back-propagation, and weight updation ( $v_{ij}$  and  $w_{jk}$ ). The feed-forward phase is the testing phase of BPN in which a number of hidden layers are used in the network to achieve the desired output. It is important to train BPN for calculation of the VMs load. There are various learning factors and activation functions that are responsible to train BPN. The use of large number of weights in BPN may slow down the convergence of the network. Hence, a Momentum Factor ( $\eta$ ) is used to save the previous information of weights for weight adjustment and for better solution. It enhances the weight updation stage and makes fast convergence. Equations (7) and (8) are the weight update expressions of the output layer units and the hidden layer units, respectively.

$$w_{jk}(t_k + 1) = w_{jk}(t_k) + \alpha(\delta_k)b_k + \eta[w_{jk}(t_k) - w_{jk}(t_k - 1)] \quad (7)$$

$$v_{ij}(t_k + 1) = v_{ij}(t_k) + \alpha(\delta_j)a_i + \eta[v_{ij}(t_k) - v_{ij}(t_k - 1)] \quad (8)$$

where  $w_{jk}$  is the output weight between  $j$ th hidden layer unit and  $k$ th output layer unit and,  $v_{ij}$  is the hidden weight between  $i$ th input layer unit and  $j$ th hidden layer unit.  $t_k$  is the targeted value of the network. To get trained output from the BPN, an activation function is used which increases monotonically. BPN mostly uses binary sigmoid function (or unipolar) that reduces computational burden during the learning process as defined in Eq. (9),

$$f(x) = 1/(1 + e^{-\lambda x}) \quad (9)$$

where  $\lambda$  is the steepness parameter.

The capacity of VM includes a number of processors, Million Instruction Per Second (MIPS) and bandwidth of that VM. The capacity and target (expected output) of BPN is calculated using the following expression:

$$C_{vmj} = (vm_{pj} \times vm_{mipsj} + vm_{bwj})/100 \quad (10)$$

$C_{vmj}$  is the capacity,  $vm_{pj}$ ,  $vm_{mipsj}$ , and  $vm_{bwj}$  are the number of processors, MIPS, and bandwidth on  $j$ th VM, respectively.  $C_{vmj}$  is used as an initial weight ( $v_{ij}$ ) on hidden layer units and calculated using Eq. (10). The initial load  $\Pi_{ij}$  on each VM denoted as  $VML = VML_1, VML_2 \dots VML_j$  is obtained by the summation of the total length of all tasks (task length as  $TL = \{TL_1, TL_2 \dots, TL_i\}$ ) on  $j$ th VM expressed in Eq. (11),

$$\Pi_{ij} = \sum_1^i TL_{ij} \in VM, i \in task \quad (11)$$

$$Eo_k = \frac{(TL_{ij}/vm_{mipsj})}{((\Pi_{ij}/vm_{mipsj})/\text{Total Number of VMs})} \quad (12)$$

where  $TL_{ij}$  is the  $i$ th task length on  $j$ th VM.  $Eo_k$  is the expected (target) load of  $j$ th VM to update the network weights through errors expressed in Eq. (12). Algorithm 1 illustrates each step of BPN that calculates the optimized load of VMs.

#### ALGORITHM 1: Back Propagation Network

- Step1:** **Start** For each VM ( $VM = VM_1, VM_2 \dots VM_j$ ), receive  $vm_{pj}$ ,  $vm_{mipsj}$ ,  $vm_{bwj}$ , and TL information.
- Step2:** Initialize input dimension, number of hidden units, number of output units, maximum epoch, learning rate ( $\alpha$ ), weight  $W_i$ , and  $T_k$ .
- Step3:** Calculate  $v_{ij}$  using Eq. (10). Set  $v_{oj} = 1$  and  $w_{ok} = 1$ . The weights on output layer  $w_{jk}$  are set as random values between 0.0 to 0.1.
- Step4:** Calculate hidden input from input layer units ( $A_i$ ),

$$h_{inj} = V_{0j} + \sum_{i=1}^n (a_i v_{ij}) \quad (13)$$

$$h_j = F(h_{inj}) \quad (14)$$

where  $h_{inj}$  refers to the hidden input signal,  $v_{0j}$  is the bias weight to hidden layer,  $a_i$  is the  $i$ th input unit,  $v_{ij}$  is the input weight from  $i$ th input unit to  $j$ th hidden unit, and  $h_i$  is the output of the  $i$ th hidden unit using Eq. (9).

- Step5:** Calculate output from the hidden layer,

$$b_{ink} = w_{ok} + \sum_{i=1}^p (h_j w_{jk}) \quad (15)$$

$$b_k = F(b_{ink}) \quad (16)$$

where  $b_{ink}$  is the output layer input signal,  $w_{0k}$  is the bias weight to output layer,  $w_{jk}$  is the input weight from  $j$ th hidden unit to  $k$ th output unit, and  $b_k$  is the output of the  $k$ th output unit calculated using Eq. (9).

**Step6:** With the target pair  $E_{ok}$  as  $T_k$  from Eq. (12), compute error-correcting factor ( $\delta_k$ ) between output layer units and hidden layer units.

$$\delta_k = (t_k - (b_k))F(b_{ink}) \quad (17)$$

Binary sigmoid activation function from Eq. (9) is used to reduce the computational burden.

**Step7:** Calculate delta output weight  $\Delta w_{jk}$  and bias correcting  $\Delta w_{0k}$  terms,

$$\Delta w_{jk} = \alpha(\delta_k h_j) \quad (18)$$

$$\Delta w_{0k} = \alpha(\delta_k) \quad (19)$$

**Step8:** Calculate error terms  $\delta_j$  between the hidden and input layer,

$$\delta_{inj} = \sum_{k=1}^m (\delta_k w_{jk}) \quad (20)$$

$$\delta_j = \delta_{inj} F(h_{inj}) \quad (21)$$

**Step9:** Calculate delta hidden weights  $\Delta v_{ij}$  and bias  $\Delta v_{0j}$  based on  $\delta_j$ ,

$$\Delta v_{ij} = \alpha \delta_j t_k \quad (22)$$

$$\Delta v_{0j} = \alpha \delta_j \quad (23)$$

**Step10:** Update output weight and bias unit using Eq. (7).

**Step11:** Update hidden weight and bias unit using Eq. (8).

**Step12:** If the specified number of epochs are reached or  $B_k = T_k$ , goto **Step13**, else goto **Step3**.

**Step13:** Calculated load of  $VM_j$ .

**Step14:** End

### 3.1.2 Improved K-Mean (iK-Mean)

The calculated load obtained from BPN forms the input to the iK-mean cluster algorithm. The algorithm uses minimum distance data points of the cluster center.



The procedure of the algorithm starts with an initial  $K$  cluster centers. The algorithm calculates the minimum distance to classify the nearest cluster center of all data points with each  $K$  selected number of centers. Next, the mean value of data to the center value is modified. The process repeats until new center value becomes equal to the previous center value. Finally,  $C_1$  (underloaded cluster) and  $C_2$  (overloaded cluster) are formed where  $C_1 = VM_U = \{VM1_U, VM2_U \dots VMi_U\}$  and  $C_2 = VM_o = \{VM1_o, VM2_o \dots VMi_o\}$ .

### 3.1.3 Particle Swarm Optimization (PSO)

Incoming user tasks are allocated to underloaded VMs to maintain load balancing. Task scheduling is performed by the PSO algorithm. PSO is a biological concept based algorithm that is inspired by flocking of birds. A swarm-based intelligence algorithm approach uses self-additive global search technique to achieve optimized results. It initializes the particle (P). These are the potential solutions that move into the problem space by subsequent current optimum particles. In PSO, each P is represented by velocity and position which are obtained using Eqs. (24) and (25), respectively. Each P adjusts its velocity and position according to its best position and the position of the best particle (global best) in the entire population at each  $k$  iteration. Fitness value (FV) is the problem specific and used to measure the performance of a particle.

$$V_P[k + 1] = w * V_P[k] + Y_1 * rand_1 * (pbest - X_P[k]) + Y_2 * rand_2 * (gbest - X_P[k]) \quad (24)$$

$$X_P[k + 1] = X_P[k] + V_P[k + 1] \quad (25)$$

where  $V_P[k + 1]$  and  $V_P[k]$  is the present velocity and earlier velocity of P, respectively.  $X_P[k + 1]$  and  $X_P[k]$  are existing and earlier locations of P. Two cognitive and social acceleration coefficients values are  $Y_1$  and  $Y_2$  respectively and  $rand_1$ ,  $rand_2$  (between 0 and 1) are used as self-regulating random numbers in the velocity computation. The population shows the particles in the search space. Best particle position in the population is denoted by  $pbest$  and  $gbest$ . The  $pbest$  is the best particle that has reached best outcome whereas best particle in global search space is denoted as  $gbest$ .  $w$  represents inertia weights that are used to balance best local and global search of particles. The value of  $w$  is a positive linear or nonlinear of time or a positive constant. Large  $w$  supports global exploration whereas small  $w$  supports local exploration. Particles are initialized randomly. The velocities synchronize P movement. At any point of time, the position of each P is influenced by its  $pbest$  in the search space. The whole working of the PSO algorithm is listed in ALGORITHM 2.

**ALGORITHM 2: Particle Swam Optimization**

```

Input: Initialize required parameters
    User tasks T= {t1, t2, t3,...,ti},k=0 and Under loaded VMU = {VM1U,VM2U,...,VMiU}
Output: Optimal mapping between T and VMU.
begin
    Set particle dimension ←number of tasks in T
    Initialize particles randomly (velocity Vi, location Xi) with pbest and gbest
    for each ti∈ T
        while (k<epoches)
            Calculate FV for each particle and update Xi
            if current FV<pbest
                Assign pbest← Xi
            else
                Keep pbest
        endif
        for each pbest
            Assign gbest← highest pbest
            if present gbest< fitness value
                Allot gbest← Xi
            else
                Keep previous gbest
        endif
        Assign VMU to particle with highest gbest
        Assign ti to VMU
        ti++
    end while
end for
end
    
```

**4 Implementation**

The proposed Artificial Neural Network based Load Balancing (ANN-LB) algorithms have been implemented for improved cloud load distribution among underloaded VMs. CloudSim tool has been used to simulate algorithms. The implementation is performed on 2.20 GHz processor with 16 GB memory.

BPN algorithm is used for load calculation of each VM. To describe the theoretical functioning of the BPN, we include five different tasks that are allocated to five different VMs. The calculation is done for one epoch set to check the leniency of the BPN algorithm. Even if VMs and tasks may increase, the working procedure will remain the same. After performing BPN algorithm, the final calculated loads of VM<sub>0</sub>, VM<sub>1</sub>, VM<sub>2</sub>, VM<sub>3</sub>, and VM<sub>4</sub> are 0.199, 0.198, 0.203, 0.199, and 0.201, respectively. VM<sub>i</sub> load obtained from BPN will form the input to iK-Mean algorithm to perform clustering between VM<sub>1</sub>, VM<sub>2</sub>,...,VM<sub>i</sub> into clusters C<sub>1</sub> (Underloaded VMs) and C<sub>2</sub> (Overloaded VMs) C<sub>1</sub> = {0.199, 0.198, 0.199} C<sub>2</sub> = {0.203, 0.201}. Once the clustered VMs are obtained, the PSO algorithm schedules the runtime tasks to the underloaded VMs. In the above-given case, initially we have five VMs and five tasks. The tasks in the dynamic time will be assigned to VMs that belong to C<sub>2</sub>. The

PSO technique initializes particles. These particles are nothing but the tasks taken from the task list that are further assigned to the VMs. Using Eqs. (24) and (25), each particle is initialized randomly with their velocity and position, respectively. A particle represents allocation of user tasks to the VMs that have available resources. This method manages load among VMs and gives load sharing facility to the cloud environment.

## 5 Result and Discussion

In the result section, experimental observations are analyzed to examine the results of ANN-LB algorithm. The proposed methods have been compared with CM-eFCFS, RR (Round Robin), MaxMin, and MinMin algorithms. The illustration of cloud metrics and achieved results is analyzed in the following section:

**MakeSpan (M)** is the completion time of VM that can be calculated from the initial scheduling procedure time up to the final task completed. In this work, M is calculated using Eq. (26).

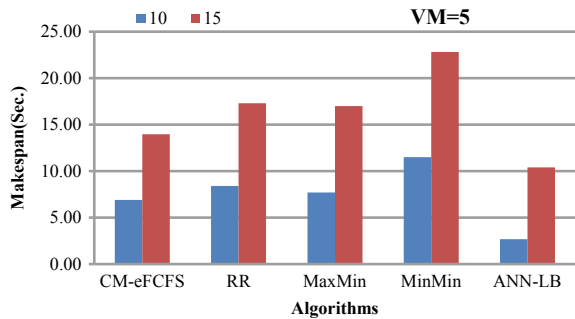
$$M = \max(CT_j) \tag{26}$$

This metric obtains the task completion time. The objective is to minimize the M to achieve faster execution. Figure 1 illustrates the total MakeSpan for CM-eFCFS, RR, MaxMin, and MinMin. Obtained results assured that introduced ANN-LB algorithm has achieved 16%, 28%, 28%, and 52% less MakeSpan than CM-eFCFS, RR, MaxMin, and MinMin, respectively.

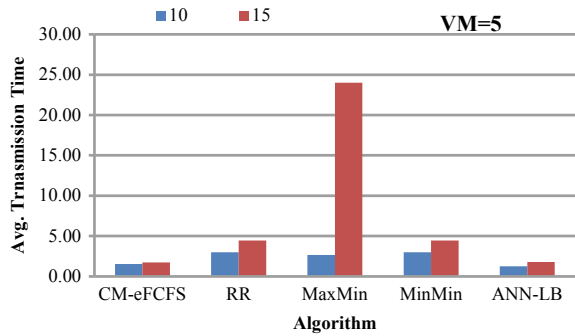
**Transmission Time (TX<sub>T</sub>)** is the time utilized to transfer the *i*th task (*t<sub>i</sub>*) on VM<sub>*j*</sub>. It is the ratio of *i*th task size and *j*th VM bandwidth. TX<sub>T</sub> is calculated using Eq. (27).

$$TX_T = \frac{Size_i}{BW_j} \tag{27}$$

**Fig. 1** Calculated MakeSpan for different algorithms



**Fig. 2** Analysis on transmission time



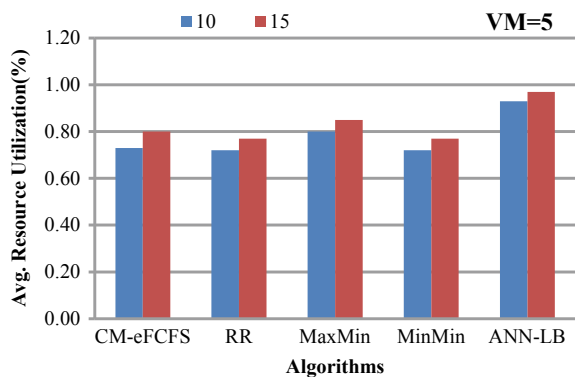
where  $Size_i$  refers to the  $i$ th task size and  $BW_j$  is the bandwidth of  $j$ th VM. Figure 2 clearly shows the graph representation of  $TX_T$  for CM-eFCFS, RR, MaxMin, and MinMin algorithms. The obtained result shows that the proposed algorithm maintains 1.16, 7.04, 1.41, and 7.04% less  $TX_T$  than CM-eFCFS, RR, MaxMin, and MinMin algorithms respectively which show better result.

**Average Resources Utilization (AU)** shows the efficiency of the system to utilize allocated cloud resources. Resource utilization should always be as high as possible to reduce wastage of resources. AU of an algorithm is evaluated using Eq. (28),

$$AU = \sum_{j \in VMs} CT_j / MakeSpan \times \text{NumberofVMs} \tag{28}$$

Figure 3 depicts the average resource utilization (AU). The obtained results clearly depict that ANN-LB algorithm attains higher utilization which is approximately 93% and 97% for 10 and 15 number of tasks, respectively. The obtained simulated results are better than CM-eFCFS, RR, MaxMin, and MinMin algorithms.

**Fig. 3** Comparative analysis on average resource utilization



## 6 Conclusion

The well responsive load-balancing technique plays the key role for a cloud computing environment that brings improvement for dynamic cloud. This paper, introduced a hybrid approach of Artificial Neural Networking and Improved K-mean clustering-based technique to achieve load balancing (ANN-LB). The role of BPN is to train the system and to get an optimized load of VMs. These optimized loads of VMs are further clustered into underloaded and overloaded. Dynamic tasks are assigned to underload VMs using PSO-based task scheduling approach. We performed the implementation in CloudSim tool and found that the effort has shown improvement of cloud metric, i.e., Resource Utilization, Transmission Time and MakeSpan. The ANN-based BPN approach has been remarkable for dynamic cloud environment. In future, we are planning to expand the proposed work for the dynamic load balancing by taking other performance parameters.

## References

1. Sosinsky, B. (2011). *Cloud computing Bible*. Wiley.
2. Shabeera, T. P., Madu Kumar, S. D., Salam, M. S., & Krishnan, K. M. (2016). Optimizing VM allocation and data placement for data-intensive applications in cloud using ACO metaheuristic algorithm. *Engineering Science and Technology, an International Journal*, 20(2), 616–628.
3. Guo, L. (2012). Task scheduling optimization in cloud computing based on heuristic algorithm. *Journal of Networks*, 7(3), 547–553.
4. Choudhary, A., Govil, M. C., Shingh, G., Aawasthi, L. K., & Pilli, E. S. (2017). A critical survey of live virtual machine migration techniques. *Journal of Cloud Computing: Advances, Systems and Application*, 6(23), 1–41.
5. Li, T., & Zhang, X. (2014). On the scheduling for adapting to dynamic changes of user task in cloud computing environment. *International Journal of Grid Distribution Computing*, 7(3), 31–40.
6. Sharkh, M. A., Shami, & Ouda, A. (2017). Optimal and suboptimal resource allocation techniques in cloud computing data centers. *Journal of Cloud Computing: Advances, Systems and Applications*. Springer Open, 6, 1–17.
7. Zhao, T., Zhou, S., Guo, X., & Niu, Z. (2017). Task scheduling and resource allocation in heterogeneous cloud for delay-bounded mobile edge computing. In *SAC Symposium Cloud Communications and Networking Track IEEE ICC*.
8. Singh, A., Juneja, D., & Malhotra, M. (2015). Autonomous agent based load balancing algorithm in cloud computing. *International Conference on Advanced Technologies and applications (ICTACTA)*, 45, 823–841.
9. Mollamotalebi, M., & Hajireza, S. (2017). Multi-objective dynamic management of virtual machines in cloud environments. *Journal of Cloud Computing: Advances, Systems and Applications*, 6(16), 1–13.
10. Hamsinezhad, E., Shahbahrami, A., Hedayati, A., Zadeh, A. K., & Baniroostam, H. (2013). Presentation methods for task migration in cloud computing by combination of yu router and post-copy. *International Journal of Computer Science Issues (IJCSI)*, 10(1), 98–102.
11. Pop, F., Dobre, C., Cristea, V., & Besis, N. (2013). Scheduling of sporadic tasks with deadline constrains in cloud environment. In *3rd IEEE International Conference on Advanced Information Networking and Application (ICAINA)*, pp. 764–771.

12. Xiao, Z., Song, W., & Chen, Q. (2013). Dynamic resource allocation using virtual machines for cloud computing environment. *IEEE Transaction on Parallel and Distributed Systems*, 24(6), 1107–1117.
13. Katyal, M., & Mishra, A. (2013). A comparative study of load balancing algorithms in cloud computing environment. *International Journal of Distributed and Cloud Computing*, 1, 5–14.
14. Li, B., Li, J., Huai, J., Wo, T., Li, Q. & Zhong, L. (2009). EnaCloud: An energy-saving application live placement approach for cloud computing environments. *International Conference on Cloud Computing. IEEE*, pp. 17–24.
15. Devi, D. C. & Rhymend Uthariaraj, V. (2016). Load balancing in cloud computing environment using improved weighted round robin algorithm for nonpreemptive dependent tasks. In *Hindawi Publishing Corporation The Scientific World Journal*, pp. 1–14.

# Maximum Power Point Tracking Using a Hybrid Fuzzy Logic Control



Amruta S. Deshpande and Sanjaykumar L. Patil

**Abstract** This paper proposes the design of a hybrid controller for a solar photovoltaic system to withdraw the maximum power. This controller combines the advantages of fuzzy controller and proportional–integral controller. The fuzzy logic control does not require the exact knowledge of the plant model while proportional–integral control reduces the offset value and gives easy architecture. Simple and efficient controller development are the main objectives behind the work. Enhanced tracking efficiency in the presence of varying environmental conditions is one of the main advantages of the controller. The controller is simulated for various irradiation and temperature conditions and the outputs are compared with fuzzy logic and traditional perturb observe controller.

**Keywords** Maximum power point tracking · Boost converter · Perturb and observe

## 1 Introduction

The demand for energy is becoming a prime challenge which can be partially solved by efficient utilization of solar energy. The photovoltaic (PV) panel tracks maximum energy when aligned properly with sunlight conditions. Maximum power point tracking (MPPT) controllers are used to overcome the low conversion efficiency problem of the solar panel. The fast convergence and low losses are the two important features of the MPPT algorithm. The comparison and review of the different power tracking algorithms help to find the power maxima of a solar module. Sensor availability, cost, selection, and application points can be considered while selecting a suitable MPPT technique [1]. Traditional perturb and observe (PO) and the incremental

---

A. S. Deshpande (✉) · S. L. Patil  
College of Engineering Pune, Pune, India  
e-mail: [asd.instru@coep.ac.in](mailto:asd.instru@coep.ac.in)

S. L. Patil  
e-mail: [slp.instru@coep.ac.in](mailto:slp.instru@coep.ac.in)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_21](https://doi.org/10.1007/978-981-15-0694-9_21)

conductance (IC) algorithm are the most popular algorithms. PO algorithm is easy to implement but the drawback of the algorithm is oscillations around a steady state which can be minimized by using variable step MPPT algorithm [2]. For finding the power maxima, irradiation, temperature, and shading conditions are very important. PV voltage and current deviation provide important information-related maximum power point (MPP) tracking or global maximum power point (GMPP) tracking [3].

The traditional tracking algorithms are easy to implement, but intelligent controllers attract most of the researchers due to their flexibility. The fuzzy logic systems (FLS) work with less accurate plant model. To implement the fuzzy controller, correct data knowledge set is required, which can be generated from an expert database. Fuzzy logic control can be effectively used for industrial applications, equipment, medical diagnostic, communications, image processing applications, automation applications, finding the maximum power of photovoltaic panel, physical tracking of the solar panel, etc. Fuzzy rule base can be developed with the knowledge of linguistic and numerical data [4]. Tracking performance of the fuzzy-based algorithm is superior to conventional algorithms in varying load conditions in various practical applications [5]. Modified hill climbing algorithm with fuzzy was developed for the PV system to achieve faster convergence speed, fewer swings around required MPP under steady state, and no difference from actual MPP for different weather conditions [6]. Fuzzy logic can be combined with a genetic algorithm (GA) to identify the MPP of the solar panel. Performance of GA optimized fuzzy control is more robust than simple fuzzy control [7].

Particle swarm optimization (PSO) can be used along with fuzzy logic to optimize membership function of fuzzy sets. It improves the transient time and MPP tracking accuracy [8]. PSO can be used in the improvement of the fuzzy-based MPPT algorithm. The fitness values of the asymmetrical fuzzy logic are higher than traditional perturb and observe and symmetrical fuzzy logic control [9]. Fuzzy based variable step-size method gives good tracking performance due to simple membership function [10]. Fuzzy logic and sliding mode control combination gives high efficiency with constant DC link voltage of a grid-connected PV system [11]. The Neural network can be effectively used to predict the different input values of the panel. The Neuro-fuzzy based algorithm estimates the climatic parameter and calculates the maximum power of a solar panel with improved power efficiency and MPPT response time [12]. The adaptation in proportional–integral–derivative (PID) control is carried out in which gain is scheduled with the fuzzy algorithm. Improved tracking is observed with a two-level fuzzy PID system under changing atmospheric conditions and different PV sources [13].

Though the combination of the different algorithms with fuzzy improves the performance of the controller, these algorithm implementations are much more complex than the simple PID controller. The proposed controller combines the advantages of fuzzy control and proportional integral called FLSPI. The development of the simple and efficient controller compared to a traditional controller is the main objective behind the work. The fuzzy control calculates maximum voltage from the knowledge of solar panel error and error change. PI controller decides the actual value



of the converter duty ratio by adjusting the gain values of the controller. Controller performance is plotted for standard, varying irradiation and temperature conditions and compared with conventional PO and fuzzy controller.

## 2 PV System

The PV system comprises the solar PV module and the dc–dc boost converter. MSX-64 SOLAREX PV module specifications are considered for the MATLAB/Simulink simulation. The solar module is made up of 36 solar cells with 64 W of maximum power ( $P_{max}$ ). The circuit diagram of the model is shown in Fig. 1. The PV cell output current  $I_{pv}$  is given by Eq. (1) which requires knowledge of photo current  $I_{ph}$ , diode current  $I_d$ .

$$I_{pv} = n_p \cdot I_{ph} - n_p \cdot I_d \quad (1)$$

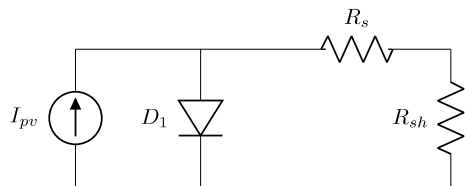
$$I_{ph} = I_{sc} + K_i [T_c - T_{ref}] \frac{G}{1000} \quad (2)$$

$$I_s = I_{rs} \left( \frac{T_c}{T_{ref}} \right)^3 \left[ \exp \left( \frac{q \cdot E_g}{kA} \left( \frac{1}{T_{ref}} - \frac{1}{T_c} \right) \right) \right] \quad (3)$$

$$I_d = I_s \left[ \exp \left( \frac{q(V_{pv})}{AkT_c n_s} \right) - 1 \right] \quad (4)$$

where  $I_{ph}$  is a photocurrent,  $I_{sc}$  is a short-circuit current,  $q$  is the charge on electron,  $k$  is the Boltzmann's constant, series and shunt resistances are defined as  $R_s$  and  $R_{sh}$ .  $K_i$  is the short circuit current temperature coefficient,  $E_g$  is the band gap energy,  $G$  is the solar irradiation,  $I_{rs}$  is the diode reverse saturation current and  $A$  is the ideality factor of a solar cell diode,  $T_c$  and  $T_{ref}$  are the surrounding temperature and reference temperature of a solar cell in Kelvin. The series and parallel cells are represented by  $n_p$  and  $n_s$ . The PV system specifications are given in Table 1. Change in irradiation and temperature condition causes the variation in I-V and P-V characteristic curve. The simulation is carried out for standard and various conditions of temperature and irradiation.

**Fig. 1** Simplified model of a solar cell



**Table 1** PV Panel and converter specification

Parameters	Symbols	Values
Power maxima	$P_{\max}$	64 W
Voltage at maximum power	$V_{\max}$	17.5 V
Current at maximum power	$I_{\max}$	3.66 A
Open-circuit voltage	$V_{oc}$	21.3 V
Short-circuit current	$I_{sc}$	4 A
Boost inductor	$L$	150 $\mu$ H
Input capacitor	$C_1$	100 $\mu$ F
Output capacitor	$C_2$	470 $\mu$ F
Resistance	$R$	50 $\Omega$
Operating frequency	$F$	50 kHz

## 2.1 Boost Converter

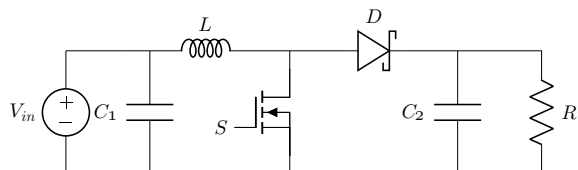
The boost converter or step-up converter is used for the simulation. The simulation parameters are shown in Table 1. Figure 2 shows the schematics of the dc–dc boost converter. Panel output is connected to the converter and the pulse width modulation (PWM) generates the appropriate pulse waveform with the calculated duty cycle to control the MOSFET switch in the dc–dc converter. The duty cycle is generated by the MPPT control algorithms.

## 3 Maximum Power Point Tracking

### 3.1 Perturb and Observe

To maximize the solar module power, voltage or current perturbation is required. If the change in voltage and power is positive, then the positive perturbation in the voltage is added and vice versa. At maximum power point (MPP), the ratio of change in power and change in voltage remains zero [1]. Maximum power is closer to MPP

**Fig. 2** Dc–dc boost converter



with smaller steps in voltage perturbation, but the major disadvantage is the time required to reach the equilibrium.

### 3.2 Fuzzy Logic

To apply effective control for the plant, several issues like the mathematical model, controller limitation, disturbances, and plant knowledge are important. Fuzzy logic system (FLS) addresses these issues by forming the rules from imprecise data. Rules can be developed with the help of computers. There are an enormous number of possibilities that lead to lots of different combinations of mapping which required a careful understanding of fuzzy logic and elements that constitute the FLS.

### 3.3 Fuzzy Logic with PI

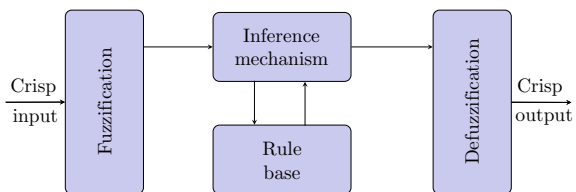
The proposed algorithm consists of the fuzzy logic system with proportional–integral (FLSPI) control for MPPT which is developed in the subsystem. Figure 3 shows the architecture of the fuzzy controller. Fuzzy controllers consist of various process stages like an input, a processing, and an output.

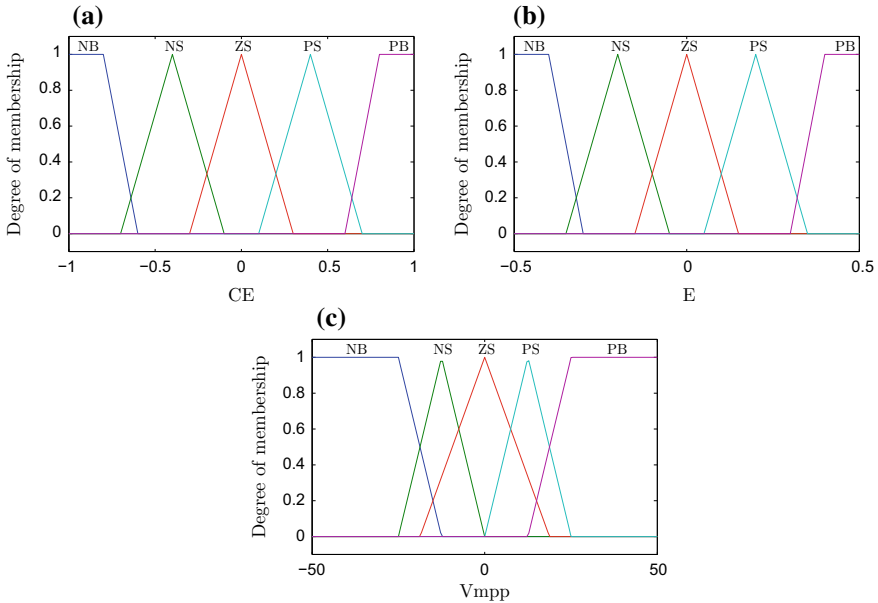
#### 3.3.1 Fuzzification

Crisp data is converted into fuzzy data by a process called fuzzification with the help of an expert’s database. For the fuzzy system, inputs are change in error and error, output is panel voltage variation. Inference mechanism involves in the decision-making. It executes various rules present in rule base to draw a conclusion. Figure 4 shows the membership function (MB) of a fuzzy set developed for the proposed control. The error E, change in error CE are the two inputs used for the fuzzy algorithm. The equations for CE and E are given below,

$$CE = e(k) - e(k - 1) \tag{5}$$

Fig. 3 Fuzzy controller





**Fig. 4** Oscillations in the PO algorithm **a** Membership for CE, **b** Membership for E, **c** Membership for panel voltage

$$E = \frac{P(k) - P(k - 1)}{V(k) - V(k - 1)} \tag{6}$$

The normalization operation is carried out on each input. Input and output consist of five membership functions: negative big, negative small, zero, positive small, and positive big represented by (NB), (NS), (ZS), (PS), and (PB), respectively. MB allows representing a fuzzy set graphically. The universe of discourse and degrees of membership [0,1] are represented on X-axis and Y-axis, respectively. Rule base consists of rules which are applied to the data given by the fuzzifier to give the output. A total of 25 rules are developed in the rule base. The example of the rules is given below. IF CE is NB and E is NB, then voltage output is PB (Table 2).

**Table 2** Rule base for duty ratio

		<i>E</i>				
		<i>NB</i>	<i>NS</i>	<i>ZS</i>	<i>PS</i>	<i>PB</i>
<i>CE</i>	<i>NB</i>	<i>PB</i>	<i>PS</i>	<i>ZS</i>	<i>NS</i>	<i>NB</i>
	<i>NS</i>	<i>PB</i>	<i>PS</i>	<i>ZS</i>	<i>NS</i>	<i>NB</i>
	<i>ZS</i>	<i>PB</i>	<i>PS</i>	<i>ZS</i>	<i>NS</i>	<i>NB</i>
	<i>PS</i>	<i>PB</i>	<i>PS</i>	<i>ZS</i>	<i>NS</i>	<i>NB</i>
	<i>PB</i>	<i>PB</i>	<i>PS</i>	<i>ZS</i>	<i>NS</i>	<i>NB</i>

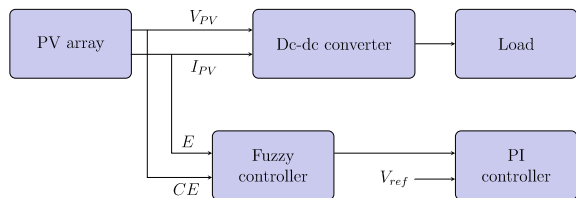
### 3.3.2 Defuzzification

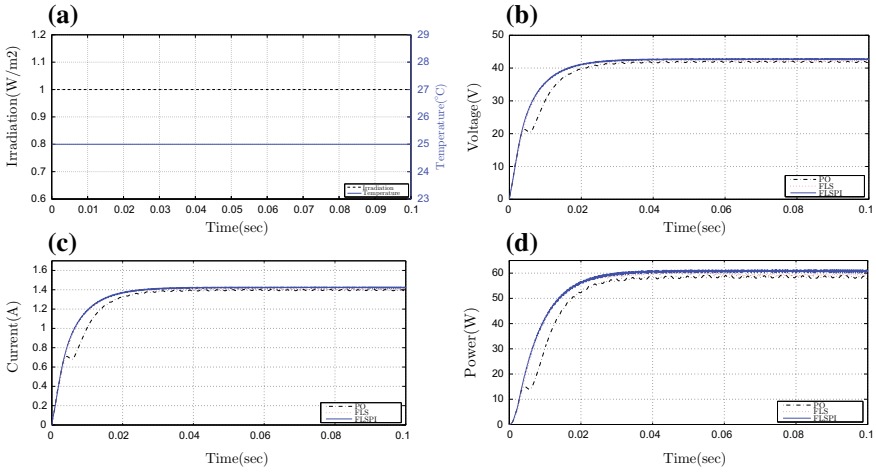
In the defuzzification process, the output signal from the membership function is converted into crisp value. Center of gravity (COG) approach is one of the simple and commonly used methods for defuzzification. The block diagram of the presented system is shown in Fig. 5. The maximum point is calculated either by calculating maximum voltage  $V_{mpp}$  or maximum current  $I_{mpp}$ . The fuzzy rule gives the output in terms of the maximum voltage. The output voltage is compared with the voltage at the peak power  $V_{mpp}$ . The input to the PI controller is the error which is the difference between the fuzzy controller output and voltage at the peak power.

## 4 Simulation Results

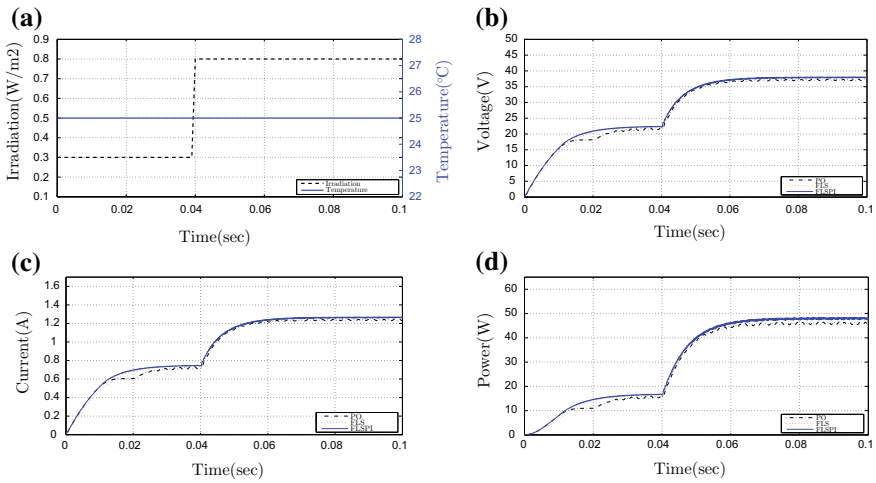
In this section, comparative study between PO, FLS, and FLSPI control algorithm is carried out. Simulink model of a PV system is developed for the series combination of 36 solar cells. Figure 6 gives voltage, current, and power output variation for the standard input condition 1 with irradiance  $1000 \text{ W/m}^2$  and a temperature of  $25 \text{ }^\circ\text{C}$ . Figure 7 gives voltage, current, and power output variation for the test scenario in which irradiation is varied from  $300$  to  $800 \text{ W/m}^2$  and temperature conditions are constant at  $25 \text{ }^\circ\text{C}$ . Figure 8 gives voltage, current, and power output variation for the test scenario in which the temperature is varied from  $25 \text{ }^\circ\text{C}$  to  $45 \text{ }^\circ\text{C}$  and irradiation conditions are constant at  $1000 \text{ W/m}^2$ . The response of the PO controller, FLS controller, and FLSPI controller is plotted with black, red, and blue color, respectively. The simulation results show that the reduction in the irradiation value reduces the output power while the increase in the temperature of the PV cell decreases the output power. Also, the response of the FLSPI system is faster than the traditional PO and FLS algorithm. The power output of the converter for the variation of the input is given in Table 3. For standard input condition, the output power (converter input) at MPP is  $64 \text{ W}$ . For PO, observed power is  $61.15 \text{ W}$  and for FLSPI, it is  $63.75 \text{ W}$ . Therefore, the efficiency of the PO algorithm is  $95.54\%$  and efficiency of the FLSPI algorithm is  $99.60\%$ .

Fig. 5 Fuzzy PI controller





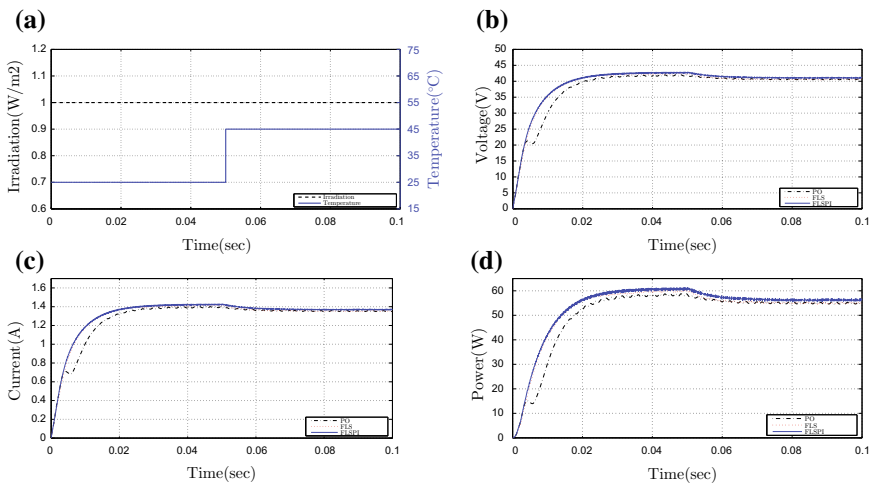
**Fig. 6** a Input condition 1 with constant irradiation and constant temperature b Variation of converter output voltage c Variation of converter output current d Variation of converter output power



**Fig. 7** a Input condition 2 with changing irradiation and constant temperature b Variation of converter output voltage c Variation of converter output current d Variation of converter output power

**Table 3** Observation table

Irradiance (W/m <sup>2</sup> )	Output power (W)		
	PO	FLS	FLSPI
1000	58.7	59.2	60.1
800	45.5	47.5	50.3
600	33.3	35.4	35.4
400	21.0	23.0	23.0
200	10.7	10.7	10.8
Temperature (°C)	Output power (W)		
	PO	FLS	FLSPI
25	58.7	59.2	60.1
35	56.5	57.6	58.4
45	54.5	55.0	55.9
55	52.5	52.1	53.7
65	50.5	51.0	51.5



**Fig. 8** a Input condition 3 with constant irradiation and changing temperature b Variation of converter output voltage c Variation of converter output current d Variation of converter output power

## 5 Conclusion

In this paper, the FLSPI controller proposed for photovoltaic system gives the improved efficiency compared to the existing (PO and FLS) algorithms. The combination of intelligent and conventional control covers the advantages of both the

methods. The results confirm the improvement in the efficiency through the solar array simulation. The proposed MPPT technique tracks the maximum power with higher (99.68%) efficiency, fast response than conventional MPPT techniques in the presence of changing environmental conditions. Auto-tuning of the membership function is the future scope of the work.

## References

1. ESRAM, T., & CHAPMAN, P. L. (2007). Comparison of photovoltaic array maximum power point tracking techniques. *IEEE Transactions on Energy Conversion*, 22(2), 439–449.
2. HOUSSAMO, I., LOCMENT, F., & SECHILARIU, M. (2013). Experimental analysis of impact of MPPT methods on energy efficiency for photovoltaic power systems. *International Journal of Electrical Power & Energy Systems*, 46, 98–107.
3. FATHABADI, H. (2016). Novel fast dynamic MPPT (maximum power point tracking) technique with the capability of very high accurate power tracking. *Energy*, 94, 466–475.
4. WANG, L. X., & MENDEL, J. M. (1992). Generating fuzzy rules by learning from examples. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(6), 1414–1427.
5. GOUNDEN, N. A., PETER, S. A., NALLANDULA, H., & KRITHIGA, S. (2009). Fuzzy logic controller with MPPT using line-commutated inverter for three-phase grid-connected photovoltaic systems. *Renewable Energy*, 34(3), 909–915.
6. ALAJMI, B. N., AHMED, K. H., FINNEY, S. J., & WILLIAMS, B. W. (2011). Fuzzy-logic-control approach of a modified hill-climbing method for maximum power point in microgrid standalone photovoltaic system. *IEEE Transactions on Power Electronics*, 26(4), 1022–1030.
7. LARBS, C., CHEIKH, S. A., OBEIDI, T., & ZERGUERRAS, A. (2009). Genetic algorithms optimized fuzzy logic control for the maximum power point tracking in photovoltaic system. *Renewable Energy*, 34(10), 2093–2100.
8. MIRHASSANI, S. M., GOLROODBARI, S. Z. M., GOLROODBARI, S. M. M., & MEKHILEF, S. (2015). An improved particle swarm optimization based maximum power point tracking strategy with variable sampling time. *International Journal of Electrical Power & Energy Systems*, 64, 761–770.
9. CHENG, P. C., PENG, B. R., LIU, Y. H., CHENG, Y. S., & HUANG, J. W. (2015). Optimization of a fuzzy-logic-control-based MPPT algorithm using the particle swarm optimization technique. *Energies*, 8(6), 5338–5360.
10. CHEN, Y. T., JHANG, Y. C., & LIANG, R. H. (2016). A fuzzy-logic based auto-scaling variable step-size MPPT method for PV systems. *Solar Energy*, 126, 53–63.
11. MENADI, A., ABDEDDAIM, S., GHAMRI, A., & BETKA, A. (2015). Implementation of fuzzy-sliding mode based control of a grid connected photovoltaic system. *ISA Transactions*, 58, 586–594.
12. CHIKH, A., & CHANDRA, A. (2015). An optimal maximum power point tracking algorithm for PV systems with climatic parameters estimation. *IEEE Transactions on Sustainable Energy*, 6(2), 644–652.
13. DOUNIS, A. I., KOFINAS, P., ALAFODIMOS, C., & TSELES, D. (2013). Adaptive fuzzy gain scheduling PID controller for maximum power point tracking of photovoltaic system. *Renewable Energy*, 60, 202–214.



# Differential Evolution Algorithm Using Enhance-Based Adaption Mutant Vector



Shailendra Pratap Singh and Deepak Kumar Singh

**Abstract** Nature-inspired optimization is the field of study for planning, simulation, and execution of problems using scientific methodologies. In this paper, a novel mutation-based modified differential evolution (DE) algorithm has been proposed. Enhance-based adaption mutation operator helps in avoiding the local optimum problem. The proposed approach is named as enhance-based adaption (EBA) in the existing mutation vector to provide more diversity for selecting effective mutant solutions. The proposed approach provides more promising solutions to guide the evolution and helps DE escaping the situation of the local optimum problem. Comparisons with other DE variants such as CPI-DE, TSDE, ToPDE, MPEDE, and JADEcr establish that the proposed Environment adaption-based operator is able to improve the performance of differential evolution algorithms.

**Keywords** Differential evolution algorithm · Enhance-based adaption factor · COCO platform

## 1 Introduction

There are various nature-inspired algorithms used to solve real-world optimization problems such as Particle Swarm Optimization (PSO) [1], Genetic Algorithm (GA) [2] and many more, but none of them guarantee to have an optimum solution better than DE. After rigorous studies and research, engineers and scientists were able to design a robust algorithm that acquires low cost and better convergence rate.

---

S. P. Singh (✉)

Department of Computer Science and Engineering, Bundelkhand Institute of Engineering and Technology Jhansi, Jhansi, UP, India  
e-mail: [shail2007singh@gmail.com](mailto:shail2007singh@gmail.com)

D. K. Singh

Department of Computer Science and Engineering, Sachedeva Institute of Technology, Frah, Mathura, UP, India  
e-mail: [yadav.k.Deepak@gmail.com](mailto:yadav.k.Deepak@gmail.com)

Evolutionary algorithm is designed to fulfill the global optimization problem [3, 4]. The anatomy of EA has been inspired by nature which conducts exploration and exploitation process through reproduction and selection operator. DE is an evolutionary algorithm, and a stochastic population-based optimization algorithm developed by Storn and Price [3, 5] and is considered to be the most recent technique to solve real-valued optimization problems [6, 7]. DE exhibits many advantages: first, it is much simpler than any other evolutionary algorithm. Its simplicity attracts researchers from other fields to solve their domain-specific problem since the main body of the program takes only four to five lines of code. Second, DE has a few parameters to work with F, NP, and Cr which influence the performance, the wrong selection of mutation operator and control parameters can lead algorithm, trapped in local optima instead of moving toward global optima. Third, overall performance is still better than other evolutionary algorithms in terms of convergence speed, accuracy, and robustness. Fourth, it requires less storage space complexity to handle large-scale and higher dimension problems.

Storn and Price [3, 5] introduced differential evolutionary (DE) as a “Simple and efficient heuristics for global optimization over continuous spaces”. In four phases, the D-dimension searches the global optimum solution in the actual parameter space, i.e., initialization, mutation, crossover, and selection. This process is explained in Algorithm 1.

---

#### Algorithm 1 Basic Differential Evolution

---

- 1: **procedure** DEA
  - 2: initialization of population  

$$\alpha_{i,G} = \{\alpha_{1,i,G}, \alpha_{2,i,G}, \dots, \alpha_{D,i,G}\}$$
  - 3: fitness value of the population
  - 4: itr = 1
  - 5: while  $FunEvs < MaxFunEvs$
  - 6: Apply generate a donor vector (mutation strategy):  

$$\gamma_{i,G} = \alpha_{best,G} + \delta_1 \cdot (\alpha_{r_1^i,G} - \alpha_{r_2^i,G}),$$
  - 7: Apply crossover generate the trail vector
  - 8: Apply the better fitness function according to the selection strategy
  - 9: itr = itr + 1
  - 10: end while
  - 11: **end procedure**
- 

Most of the articles on differential evolution [8–14] discussed problems of stagnation. These algorithms are either single objective or multi-objective to provide the solution for the problem of stagnation. The provided solution by these algorithms are based on Homeostasis and adaption-based mutation and increase the convergence speed. These algorithms also performed better than other optimization algorithm on standard benchmark functions. In this paper, a new “Differential Evolution algorithm using Enhance-based adaption mutant vector (EABMO-DEA)” has been discussed. EABMO-DEA is a new variant of DE that is introduced to solutions from merging to locally optimal solutions and also to maintain the diversity and increase the

convergence rate of the algorithm. The algorithm is tested for performance by comparing it with other sophisticated algorithms. Analysis conducted on performance posits that EABMO-DEA algorithm performs better than other modern DE algorithms.

The evaluation of the performance of EABMO-DEA has been conducted on CEC-2015 benchmark functions and it has been observed that EABMO-DEA better performs than other optimization algorithms on standard benchmark functions. The rest of this paper is organized as follows: in Sect. 2, the proposed approach Enhance-based adaption based operator is explained, Sect. 3 describes experimental result and analysis by comparing variants of DE, and in Sect. 4, conclusion and future work of this paper are described.

## 2 Proposed Approach: DE Algorithm Using Enhance-Based Adaption Mutant Vector

Improvements introduced in the performance of DE variants are actually dependent on control parameters, reproduction operator, and selection procedure. These operators use the exploitation and exploration procedure of the global search around the candidate of population. In the same way, an Enhanced-based adaptation technique is designed by multiplying with each corresponding target vector, thus, retaining the diversity from the initial generations till the end. The designed strategy multiplies these vectors to derive more area for exploration while others make use of it for exploitation. This improved mutation strategy is named after their basic mutation strategy which is the common procedure for creating new enhance-based adaption variant of DE. The algorithm improves the convergence speed, maintains the diversity of global search, and also reduces the fitness evaluations. The convergence speed is improved due to introduced better candidate in the search procedure exploration and exploitation so that it does not stagnate in mutation.

An Enhance-based adaption (EBA) means finding and maintaining better solutions in the search space. Equations (1) and (2) sustain the internal system of the search space through the designed adaptation technique. These equations maintain (adapt the environment of the population) the environment of search space in a given area:

$$EAB1 = random\_vector * EAF \quad (1)$$

$$EAB2 = current\_vector * EAF \quad (2)$$

where *current\_vector* represents the current population of global search space and *random\_vector* represents the given population in the search area. An enhance-based adaptation factor (EAF) is a persistent adaptation variation which works

according to the designed fitness function. For this proposal, the chosen value of EAF is 0.1–1.2.

The EBAs maintain the diversity and convergence rate when it is stuck in a local optimum problem, then the sufficient requirement to provide diversity is achieved by introducing EAF value to (0.1, 1.2).

To generate Enhance-based adaption based mutation operator of DE/best/1, Enhance-based adaption are multiplied by an the existing strategy. First, a random difference vector is created and then the Enhanced adaption (EA) is multiplied to the basic mutations strategy in mutation operators. We define the general mutation operator in Eq. (3), and multiply the enhance adaption factor (EAF)-based mutation operator in Eq. (4).

$$\gamma_{i,G} = \alpha_{\text{best},G} + \delta_1 \cdot (\alpha_{r_1^i,G} - \alpha_{r_2^i,G}) \quad (3)$$

Generate the new mutant vector:

$$\gamma_{i,G} = \alpha_{\text{best},G} + \delta_1 \cdot (\mathbf{EAB}_{r_1^i,G} - \mathbf{EAB}_{r_2^i,G}) \quad (4)$$

where  $\alpha_{\text{best}}$  denotes the best vector of the current population.  $\mathbf{EAB}_{r_1^i,G}$  and  $\mathbf{EAB}_{r_2^i,G}$  will be generated from the entire search space. This process is explained in Algorithm 2.

---

### Algorithm 2 EABMO-DEA

---

- 1: Population based initial values of parameters as
    - $\delta_1 = \text{rand}/3$ , where rand value (0 to 1)
    - $Cr = 0.4$ ;
    - $AF = 0.1$  to 1;
  - 2: PopSize = 50\*D
  - 3: Initialize of fitness function of population (randomly)
  - 4: itr = 1
  - 5: while  $FES < MaxFES$
  - 6: Apply Enhance-based adaption based mutation operator and generate donor vector:
 
$$\gamma_{i,G} = \alpha_{\text{best},G} + \delta_1 \cdot (\mathbf{EAB}_{r_1^i,G} - \mathbf{EAB}_{r_2^i,G})$$
  - 7: Apply crossover operator and generate trail vector using eq (2)
  - 8: Apply Selection operator
  - 9: itr = itr + 1
  - 10: end while
- 

## 3 Experiment and Analysis

The proposed method of EABMO-DEA has been evaluated using Black-Box Optimization Benchmarking (BBOB) in the COCO platform for the 24 [1, 4, 15, 16] noiseless test functions. These functions are based on BBOB, which is five groups in the benchmark function are given below:

1. Group1: (f1–f5) is the first benchmark function that is called the separable Functions.
2. Group2: (f6–f9) is the second benchmark function that is called the moderate conditioning.
3. Group3: (f10–f14) is the third benchmark function that is called the high conditioning.
4. Group4: (f15–f19) is the fourth benchmark function that is called the adequate global structure.
5. Group5: (f20–f24) is the fifth benchmark function that is called the multimodal functions.

### 3.1 Environment of Testing Framework for Benchmark Functions

EABMO-DEA is tested using the COCO framework. The area of search space is  $[5, -5]^D$ , that is, the vector can have a mutant and crossover- based vector within the domain  $[-5, 5]$ . Most of the benchmark functions have optima in the domain  $[-4, 4]$ . The test is done to reduce the problem. Function evaluation of 15 time instances of each ceremony is taken. Termination conditions are either more than  $10^{-8}$  functionality evaluation or precision. EABMO-DEA uses different standards, whose value is explained as the population size which is taken as 50. The control parameter is explained in Algorithm 2. The table represents the symbols for the proposed algorithm and environment of the Testing Framework. The environment of the Testing Framework used the minimum and maximum number of functional evaluations (*FES*). The function that applies to the benchmark functions of the algorithm is  $10000 * D$ , where D is the dimension like 10-D of different search space.

The environment experiments were done on a personal computer with Intel Core i7-8850H Processor i7 CPU with a speed of 2.60 GHz, and RAM (memory) 8 GB, operating system Windows 10 Pro 64 bit and x64- based processor.

### 3.2 Result Analysis

The proposed method is compiled and compared with standard DE algorithms like JADEcr [17], ToPDE [18], MPEDE [19], TSDE [20], and CPI-DE [21] on 10 dimensions (10D), which justifies the rationale for the improvement of the convergence speed for all the work result analysis (f1–f24). This function is used to test against the standard target function  $10^{-8}$  and the result of the proposed method is found to be better than the other existing modern algorithms in terms of the minimum number of function evaluations. Statistical representation of functions including “Function Evaluation (FES)” and “Best Search” are shown in Table 1.





### 3.3 *Effect of Various Parameters on the Performance of the Algorithm*

NP (Population size) is an important parameter for any DE algorithm. The size of the population is dependent on the introduced dimensions, that is, smaller to higher; these dimensions are used in capturing the target value for a given function. This is due to the design of the adaptation operator. If we take a large size, we may do exploration and exploitation around many local areas. The number of generation is important for a DE algorithm. For this proposal, 50 number of generations may work efficiently with any number of populations.

The function of mutant factor ( $\delta 1$ ) and crossover  $Cr$  are fixed, but it is tuned if we are not getting better results according to the optimization problem. Thus, we obtain better candidate solutions that improve the convergence speed and maintain diversity.

## 4 Conclusion

In this paper, enhance-based adaption mutation operator using differential evolution strategy has been seen to have enhanced performance on the 10 dimensions. Enhance-based adaption (EBA) helps in avoiding the local optimum problem. This problem is solved by taking the value of EAF from 0.1 to 1.2; this value is chosen according to the fitness function to improve the convergence rate in the search space. The comparison with the existing standard DE algorithms such that CPI-DE, JADEcr, TSDE, MPEDE, and ToPDE on 10 dimensions (10D) have been made and it shows enhanced performance in the context of EABMO-DEA.

The proposed approach performed almost very well in all kinds of functions (f1–f24). This proposed algorithm with very low downfall has been seen in very high dimensions. This algorithm is used as a future work to enhance the dynamics of EAF. The concept of the improved dynamics of EAF will be effectively applied in order to improve the performance of the proposed algorithm on higher dimensions.

## References

1. <http://coco.gforge.inria.fr/>.
2. Andre, J., Siarry, P., & Dognon, T. (2001). An improvement of the standard genetic algorithm fighting premature convergence in continuous optimization. *Advance in Engineering Software*, 32, 49–60.
3. Das, S., & Suganthan, P. N. (2011). Differential evolution: A survey of the state-of-the-art evolutionary computation. *IEEE Transactions on Evolutionary Computation*, 15(1), 4–31.
4. Brown, C., Jin, Y., Leach, M., & Hodgson, M. (2015, June 27).  $\mu$ JADE: Adaptive differential evolution with a small population. *Soft Computing*.



5. Rainer, S., & Kenneth. P. (1995). *Differential evolution: A simple and efficient adaptive scheme for global optimization over continuous spaces*. Berkeley, CA: International Computer Science Institute.
6. Zaharie, D. (2002). Critical values for the control parameters of differential evolution algorithms. In *Proceedings of the 8th International Mendel Conference on Soft Computing* (pp. 62–67).
7. Zhang, J., & Sanderson, A. C. (2009). JADE: Adaptive differential evolution with optional external archive. *IEEE Transactions on Evolutionary Computation*, 13(5), 945–958.
8. Singh, S. P., & Kumar, A. (2018). Multiobjective differential evolution using homeostasis based mutation for application in software cost estimation. *Applied Intelligence*, 48(3), 628–650.
9. Singh, S. P., Kumar, A. (2017). Software cost estimation using homeostasis mutation based differential evolution. In *2017 11th International Conference Intelligent Systems and Control (ISCO)*, (pp. 171–181).
10. Singh, S. P., Kumar, A. (2017). Homeostasis mutation based differential evolution algorithm. *Journal of Intelligent & Fuzzy Systems*, 32(5), pp. 3525–3537.
11. Singh, S. P., Kumar, A. (2017). Pareto based differential evolution with homeostasis based mutation. *Journal of Intelligent & Fuzzy Systems*, 32(5), 3245–3257.
12. Singh, S. P., Singh, V. P., Mehta, A. K. (2018). Differential evolution using homeostasis adaption based mutation operator and its application for software cost estimation. *Journal of King Saud University-Computer and Information Sciences*.
13. Singh, S. P., & Kumar, A. (2017). Differential evolution algorithm using population-based homeostasis difference vector. *Advances in Computer and Computational Sciences*, 579–587. Springer's.
14. Singh, S. P. (2019). New adaption based mutation operator on differential evolution algorithm. *Intelligent Decision Technologies*, in press.
15. Holtschulte, N., & Moses, M. (2013). Benchmarking cellular genetic algorithms on the BBOB noiseless testbed. In *GECCO13 Companion*, Amsterdam, Netherlands.
16. Tanabe, R., & Fukunaga, A. (2015). Tuning differential evolution for cheap, medium, and expensive computational budgets. In *IEEE Congress on Evolutionary Computation (CEC)*, Press.
17. Gong, W., Cai, Z., & Wang, Y. (2014). *Repairing the crossover rate in adaptive differential evolution* (pp. 149–168). Elsevier: Applied Soft Computing.
18. Wang, Y., Xu, B., Sun, G., & Yang, S. (in Press). A two-phase differential evolution for uniform designs in constrained experimental domains. *IEEE Transactions on Evolutionary Computation*, <https://doi.org/10.1109/TEVC.2017.2669098>.
19. Wu, G. H., Mallipeddi, R., Suganthan, P. N., Wang, R., Chen, H. K. (2016). Differential evolution with multi-population based ensemble of mutation strategies. *Information Sciences*, 329-345, <https://doi.org/10.1016/j.ins.2015.09.009>.
20. Liu, Z. Z., Wang, Y., Shengxiang, Y., & Cai, Z. (2016). Differential evolution with a two-stage optimization mechanism for numerical optimization. In *2016 IEEE Congress on Evolutionary Computation (CEC)* (pp. 3170–3177). Vancouver, BC.
21. Wang, Y., Liu, Z. Z., Li, J., Li, H. X., & Yen, G. G. (2016). Utilizing cumulative population distribution information in differential evolution. *Applied Soft Computing*, 48, 329–346.

# Standard Library Tool Set for Rough Set Theory on FPGA



Vanita Agarwal and Rajendrakumar A. Patil

**Abstract** Rough Set Theory is a powerful Artificial Intelligence based tool used for data analysis and mining Inconsistent Information Systems. In the presence of inconsistent, incomplete, imprecise or vague data, normal statistical-based data analytic techniques lag behind. The various software used for the analysis of inconsistent data using Rough Set Theory runs on x86 kind of processors for various operating systems. Unlike the other software implementations, the main objective of undertaking this experimentation is to describe a new and standard library tool set for the computation of inconsistent data using Rough Set Theory which is completely synthesizable on FPGA. Further, the authors have also studied the effect of implemented design on Zybo FPGA for understanding the area, timing, and power efficiency criteria. A Rough Set Theory based Data Analytic Engine can be used as a potential candidate for knowledge discovery and data mining of inconsistent data in IoT applications at fog and/or edge interfaces. This paper defines the standard library tool for Rough Set Theory on FPGA.

**Keywords** Rough set theory (RST) · Internet of things (IoT) · Inconsistent information systems (IIS) · Field programmable gate array (FPGA)

## 1 Introduction

An Inconsistent Information System signifies inconsistent, incomplete, vague and/or imprecise data. In 1982, Prof. Pawlak Z. proposed Rough Set Theory (RST) [1–3] for reasoning/mining inconsistent data by evaluating equivalence relations between two sets of inconsistent data and partitioning it on the basis of concepts generated.

---

V. Agarwal (✉) · R. A. Patil  
Electronics and Telecommunication Department, College of Engineering Pune, Pune,  
Maharashtra, India  
e-mail: [vsa.extc@coep.ac.in](mailto:vsa.extc@coep.ac.in)

R. A. Patil  
e-mail: [rap.extc@coep.ac.in](mailto:rap.extc@coep.ac.in)

Rough Sets can perform better than statistical analysis when the underlying data distribution that deviates significantly from a normal distribution is inconsistent, or in cases where the sample size is very small [4]. Rough Set can also be used for data-limited Machine learning techniques, i.e., it needs a small number of training examples. It does not require lots of data sets to train, unlike the neural networks. Rough Sets can be used for large scope computing, scientific, and financial analysis, Fault tolerability (Missing Attributes [5]), etc. Several researchers have explored usage of Rough Set Theory for various applications [6–14].

The ultimate aim is to design a Data Analytic Engine using Rough Set Theory for Internet of Things (IoT) Applications at fog and/or edge interfaces [15]. Toward the software development for the adaptation in the design of Data Analytic Engine, the authors in this paper describe a new and standard library tool set for Rough Set Theory which is completely synthesizable on FPGA. For achieving complete synthesis, the authors have avoided several data structure constructs like linked lists which are pure software constructs and cannot be synthesized on FPGA. For authenticating their results, authors have used their own new library tool set and compared the output using the ROSE software for analyzing inconsistent data. A small inconsistent data set and its analysis is discussed in further sections comparing the output generated using the proposed new library tool and the ROSE software. The authors have also studied the effect of the implemented design on Xilinx Zybo FPGA for understanding the area, timing, and power efficiency criteria.

The remainder of this paper is organized as follows. Section 2 highlights the various software and hardware implementations for Rough Set Theory. Section 3 discusses various Rough Set Theory Constructs used for mining inconsistent data. Section 4 discusses the experimentation done for defining the standard library tool set for RST and its adaptation in the design of Data Analytic Engine. Section 5 highlights the results obtained using the ROSE Software matches with the outputs obtained through Vivado HLS 2017.2. In the end, the paper concludes by defining the standard library tool for Rough Set Theory on FPGA.

## 2 Related Work

For computation purposes of Rough Set Theory (RST), various software such as ROSE [16], RSES, WEKA, ROSETTA, Rough Sets [9, 17] are available which can be used. While ROSE and ROSETTA can run only on Windows platforms, Rough Sets, RSES, and WEKA can run on Windows/Linux/Mac platforms. These software are not completely synthesizable on FPGA because they use a linked list and other data structures in their construction of the library set. Software approach provides flexibility; however, it becomes slow when a large amount of data (trillions of data for IoT) needs to be handled during processing due to storage of raw data represented as information tables. Hardware implementation can provide efficient handling of a large amount of data.

In the past, several researchers have discussed the concept of Rough Set processor initiated by Pawlak [19], and then that of Muraszkievicz and Rybinski [20]. Later on, Lewis et al. [21] explored the self-learning hardware model based on RST Constructs using the Xilinx board.

Kanasugi [22], Kanasugi and Matsumoto [23] and Kanchan et al. [18, 25] tried designing the Rough Set coprocessor by computing the discernibility matrix. Sun et al. [24] tried using the genetic algorithm-based attribute reduction system.

Maciej Kopczyński et al. [26–28] computed short reduct and core based on discernibility matrix on FPGA.

Most of the abovementioned hardware implementations involve the sharing of the main memory of the computer. This kind of implementation in real time can cause huge computational workloads on the main machines due to the inconsistent nature of data in an IoT environment at fog and/or edge interfaces. The authors propose to offload the main machine for data analytic by adding a separate coprocessor or a hardware accelerator [15]. For experimentation, Xilinx Zybo FPGA kit has been selected due to its rich features.

### 3 Rough Set Theory to Analyze Inconsistent Information Systems

An information system consists of the universe  $U$ , condition attributes (simple values), and decision attributes (decision).  $U$  is a universal, non-empty finite set representing the objects/instances/cases. The attribute set is a non-empty finite set representing the attributes/features. The rows represent objects while the columns represent attribute values belonging to these objects [29]. Condition attributes are independent variables and decision attribute is a dependent variable and is denoted by  $d$ . A representative Inconsistent Information System case study is taken in Table 1 showing Patient Diagnostic System in an IoT environment. **An information system is called inconsistent when the same condition attribute values lead to different concepts.**

**Table 1** Inconsistent information table showing patient diagnostic system

Patient	Blood sample	Muscle pain	Temperature	Malaria
P1	Yes	Yes	Normal	Yes
P2	Yes	Yes	Very high	No
P3	Yes	Yes	High	Yes
P4	Yes	Yes	Normal	No
P5	No	No	High	No
P6	No	Yes	Very high	Yes

Table 1 contains universe  $U$  of elements, e.g., Patient.

Attributes = Blood Sample, Muscle Pain, Temperature, Malaria, etc.  
 Condition attributes = Blood Sample, Muscle Pain, and Temperature  
 Decision attribute = Malaria

The subsections given below show various constructs for computation of rules using RST.

### 3.1 *Elementary and Crisp Set*

Equivalence relations can be defined on Condition and Decision attributes. Elementary Set is defined by the equivalence relations on Conditional Attributes and Crisp Set is defined by equivalence relations on Decision Attributes. Equivalence relation in a set satisfies reflexive, symmetric, and transitive properties.

### 3.2 *Approximation*

In the algebraic space, Rough Set Theory approximates the given concept(s) using lower and upper sets of the concept(s)

1. Lower approximation and positive region ( $L_A$ ): The elements that certainly belong to the set.
2. Upper approximation ( $U_A$ ): The elements that possibly belong to the set.

The boundary region: Boundary region =  $U_A - L_A$ .

### 3.3 *Accuracy of Approximation*

It is the measure of how closely the Rough Set is approximating the target set. Accuracy measure  $\alpha R(X)$ : the degree of completeness of our knowledge  $R$  about the set  $X$  [29].  $\alpha R(X) = U_A/L_A$ .

### 3.4 *Rule Generation*

During Rule generation, the decision rules that are minimal and yet describe the data accurately are obtained.

### 3.5 *Reduct and Core*

A reduct is a set of sufficient condition attributes equivalent to the original data. Rules can be derived from reduct. The common reduct is defined as the core.

## 4 Experimentation

Authors in this paper have defined the standard library tool set (source code) for Rough Set Theory for adaptation in the design of Data Analytic Engine. Python/C can be the language of choice as the definition of library tool is non-processor-specific for IoT environment, and can be used for any hand-held portable gadgets.

The library (source code) for RST is defined using C language. Various constructs of RST such as elementary set, crisp set, lower approximation, upper approximation, core and reduct are designed using C. Our C code is different from others so that it can be completely implemented on hardware. **Data structures such as linked list and others were completely avoided.**

The code was compiled and debugged, and the output was generated using GCC compiler. The outputs match with the ROSE software outputs. The codes were also run on DSP and embedded hardware platforms like DSP TMS320 C6713 and ARM Cortex M4 boards, respectively. Code profiling was also performed on both the boards [30]. The Code profiling results showed that many NOP operations were executed and the processor was stalled for a long time, remaining idle. So we could conclude that specific instruction set for the support of Rough Set Theory is required.

The adaptation of the library tool set is done through simulation and synthesizing from C to Verilog using Vivado HLS 2017.2 and Vivado 2017.2. Vivado HLS maps the Elementary Set Source Code to the corresponding logic module. The body of the function specifies the calculation of the Elementary Set values from the input values in each function invocation. A relation between Blood Sample and Temperature as shown in Table 1 is considered to generate the concept value. This simple example maps to a control path with six states and the data flow module with no side effects in the data flow between the input and output.

## 5 Results and Discussion

**The proposed work is divided into two parts. The first part defines the standard library tool set (source code) for Rough Set Theory and the second part adapts the source code in the design of Data Analytic Engine. The first part is justified by looking at the ROSE software outputs and C source code compiled output.** Tables 2, 3, and 4 show the approximation result, Core result, and Rule induced using the ROSE software. Tables 5, 6, and 7 show the Approximation results and Table 8

**Table 2** Lower and upper approximation generated using rose software

Class	No of objects	$L_A$	$U_A$	Accuracy
N	3	2	4	0.5000
Y	3	2	4	0.5000

**Table 3** Core generated using rose software

Decision	Core
D1	Bloodsample
D1	Temperature

**Table 4** Rules generation using rose software

Rules	Relation	Similarity
Rule 1	$(\text{Musclepain} = 0) \Rightarrow (\text{Malaria} = 0)$	[P5]
Rule 2	$(\text{Bloodsample} = 1) \ \& \ (\text{Temperature} = 2) \Rightarrow (\text{Malaria} = 0)$	[P2]
Rule 3	$(\text{Bloodsample} = 0) \ \& \ (\text{Musclepain} = 1) \Rightarrow (\text{Malaria} = 1)$	[P6]
Rule 4	$(\text{Bloodsample} = 1) \ \& \ (\text{Temperature} = 1) \Rightarrow (\text{Malaria} = 1)$	[P3]
Rule 5	$(\text{Temperature} = 0) \Rightarrow (\text{Malaria} = 0) \ \text{OR} \ (\text{Malaria} = 1)$	[P4, P1]

**Table 5** Approximation result with relation blood sample and temperature using vivado HLS 2017.2

Class	$L_A$	$U_A$
Y	[P3, P6]	[P1, P3, P4, P6]
N	[P2, P5]	[P1, P2, P4, P5]

shows the Core result generated using Vivado HLS 2017.2 which has the inbuilt GCC Compiler. The Approximation result in Table 5 shows the Approximation for one relation between Blood Sample and Temperature. Approximation results from other relations are also tabulated. In Table 8, B represents Blood Sample and T represents Temperature.

The ROSE software generated results as shown in Tables 2, 3, and 4 match with the Vivado HLS generated result as shown in Tables 5, 6, 7, and 8. Thus our implementation of standard library tool set is justified.

Various rules that can be generated from Lower and Upper Approximation under the relation Blood Sample and Temperature using the proposed library set are as follows:

1. If Blood Sample is Yes (Bloodsample = 1) and Temperature is High (Temperature = 1), then Malaria is Yes (Malaria = 1)—Rule 4 in Table 4 for Patient P3.

**Table 6** Approximation result with relation blood sample and muscle Pain using vivado HLS 2017.2

Class	$L_A$	$U_A$
Y	[P6]	[P1, P2, P3, P4, P6]
N	[P5]	[P1, P2, P3, P4, P5]

**Table 7** Approximation result with relation muscle pain and temperature using vivado HLS 2017.2

Class	$L_A$	$U_A$
Y	[P3]	[P1, P2, P3, P4, P6]
N	[P5]	[P1, P2, P4, P5, P6]

**Table 8** Core generated using vivado HLS 2017.2

Decision	Core
D1	Bloodsample
D1	Temperature

2. If Blood Sample is Yes (Bloodsample = 1) and Temperature is V. High (Temperature = 2), then Malaria is No (Malaria = 0)—Rule 2 in Table 4 for Patient P2.

Rules inducted from other relations like Blood Sample and Muscle pain or Muscle pain and Temperature using the proposed library set are as follows:

1. If Blood Sample is No (Bloodsample = 0) and Muscle pain is Yes (Musclepain = 1), then Malaria is Yes (Malaria = 1)—Rule 3 in Table 4 for Patient P6.
2. If Muscle pain is No (Musclepain = 0), then Malaria is No (Malaria = 0)—Rule 1 in Table 4 for Patient P5.
3. If Temperature is Normal (Temperature = 0), then Malaria can be Yes (Malaria = 1) or Malaria can be No (Malaria = 0)—Rule 5 in Table 4 for Patient P4 and P1.

**The second part of the proposed work is justified by observing the synthesized and implemented timing and utilization reports.** Table 9 tabulates the synthesis report when xc7z010clg400-1 (Zybo) FPGA kit was used. Table 10 tabulates the synthesis and implementation utilization reports for the above implementation. Tables 9 and 10 show the timing and area utilization report for the proposed library set. The report also shows that the implemented design requires less hardware than the synthesized design. From Table 9, it is also evident that DSP48E has not been utilized. Therefore MAC operations are not a part of RST calculations. We obtained similar results during code profiling on DSP TMS320 C6713 where the Multiplier unit was never used [30].



**Table 9** Vivado HLS 2017.2 Synthesis report generated using xc7z010clg400-1 (Zybo) FPGA

Synthesized parameters	Elementary and crisp set	$L_A - U_A$	Reduct, boundary, accuracy	Core
Estimated clock period (ns)	6.56	5.32	0	4.74
Worst case latency (clock Cycles)	1800	1711	0	15
No. of DSP48E used	0	0	0	0
No. of FFs used	549	639	0	66
No. of LUTs used	682	1155	0	143

**Table 10** Vivado 2017.2 synthesis and implementation utilization report generated

RST constructs	Synthesized design		Implemented design	
	LUT	FF	LUT	FF
Elementary and crisp set	661	524	472	524
$L_A$ and $U_A$	551	503	308	503
Reduct, boundary, accuracy	0	0	0	0
Core	11	13	11	13

These timing and area utilization report can only be optimized for the proposed library set for Rough Set Theory, but there is not much scope for optimizing hardware by following this procedure (synthesizing the C program). For the best and worst-case analysis, authors suggest hardware-based design and optimization for the Data Analytic Engine.

## 6 Conclusion

Rough Set theory can be used as a novel data mining technique for analyzing inconsistent, incomplete, imprecise or vague nature of Big Data. In this paper, the authors have proposed a new and standard library tool set (source code) for Rough Set Theory for the computation of inconsistent data in IoT environment at fog and/or edge interface. The proposed tool set is completely synthesizable on FPGA. The proposed library set has been defined for adaptation in the design of Data Analytic Engine/Coprocessor/hardware accelerator. The outputs obtained by synthesizing the proposed library set for Rough Set Theory using Vivado HLS 2017.2 is authenticated with the outputs obtained using the ROSE Software. From the implemented area utilization reports and timing reports, authors suggest for looking at hardware-based design and optimization for the best and worst-case analysis as future scope of work.

**Acknowledgements** The authors would like to thank Mr. A. B. Patki, Ex-Senior Director/Scientist G and HoD, Ministry of Electronics and Information Technology, Government of India for his valuable suggestions and guidance. We also acknowledge the help provided by the officials of the College of Engineering Pune.

## References

1. Pawlak, Z. (1982). Rough sets. *International Journal of Computer and Information Science*, 11, 341–356.
2. Pawlak, Z. (1984). Rough classifications. *International Journal of Man Machine Studies*, 20.
3. Pawlak, Z. (1991). *Rough sets: Theoretical aspects and reasoning about data*. Kluwer Academic.
4. Araujo, R., & Borges, M. (2001). Extending the software process culture—an approach based on groupware and workflow. In F. Bomarius & S. Komi-Sirviö (Eds.), *PROFES 2001* (Vol. 2188, pp. 297–311)., LNCS Heidelberg: Springer. [https://doi.org/10.1007/3-540-44813-6\\_26](https://doi.org/10.1007/3-540-44813-6_26).
5. Jiye, L., & Cercone, N. (2006). Assigning missing attribute values based on rough sets theory. *Proceedings of IEEE International Conference on Granular Computing, GrC, 2006* (May), 10–12.
6. Verma, N., et.al. (2011). Rough set techniques for 24 hour knowledge factory. In *Proceedings of the 5th National Conference; INDIACOM-2011 Computing For Nation Development*. Retrieved March 10–11, 2011.
7. Patki, A. B., & Verma, S. (2009). Implementing data mining software modules using rough set techniques. In *Proceedings of National Conference on Recent Developments in Computing and its Applications, NCRDCA.09*. New Delhi: Department of Computer Science, JamiaHamdard. Retrieved August 12–13, 2009.
8. Patki, T., Kapoor, A., Khurana, S. (2005). Analytical methodologies in soft computing: Rough sets techniques. *Training Report No. DIT/D(ABP)/MSIT/05*, July 2005.
9. Riza, L. S., et al. (2014). Implementing algorithms of rough set theory and fuzzy rough set theory in the R package “RoughSets”. *Information Sciences*, 287, 68–89.
10. Hassan, Y. F. (2017). Deep learning architecture using rough sets and rough neural networks. *Kybernetes*, 46(4), 693–705. <https://doi.org/10.1108/K-09-2016-0228>.
11. Zhang, Q., Xie, Q., & Wang, G. (2016). A survey on rough set theory and its applications. *CAAI Transactions on Intelligence Technology*, 1, 323–333.
12. Bello, R., & Falcon, R. (2017). “Rough sets in machine learning: A review”, chapter in studies in computational. *Intelligence*. [https://doi.org/10.1007/978-3-319-54966-8\\_5](https://doi.org/10.1007/978-3-319-54966-8_5).
13. Hua, J. Study on the application of rough sets theory in machine learning. In *Proceedings of Second International Symposium on Intelligent Information Technology Application*. <https://doi.org/10.1109/IITA.2008.154>
14. Mahajan, P., Kandwal, R., & Vijay, R. (2012). Rough set approach in machine learning: A review. *International Journal of Computer Applications* (0975-8887), 56(10), October 2012.
15. Agarwal, V., Patil, R. A., & Patki, A. B. Architectural considerations for next generation IoT processors. *Accepted for Publication in IEEE Systems Journal*. <https://doi.org/10.1109/JSYST.2018.2890571>
16. ROSE 2 User guide. (2017). Retrieved June 25th, 2017, from <http://idss.cs.put.poznan.pl/site/fileadmin/projects-images/rosemanual.pdf>.
17. Abbas, Z., & Burney, A. (2016). A survey of software packages used for rough set analysis. *Journal of Computer and Communications*, 4, 10–18.
18. Tiwari, K. S., Kothari, A. (2014). Design and implementation of rough set algorithms on FPGA: A survey. *International Journal of Advanced Research in Artificial Intelligence*, 3(9).
19. Pawlak, Z. (2004). Elementary rough set granules: Toward a rough set processor. *Rough-Neural Computing Cognitive Technologies*, 5–13.
20. Muraszkievicz, M., & Rybinski, H. (1994). Towards a parallel rough set computer. *Springer: Rough sets, fuzzy sets and knowledge discovery* (pp. 434–443).
21. Lewis, T., Perkowski, M., & Jozwiak, L. (1999). Learning in hardware: Architecture and implementation of an FPGA—Based rough set machine. In *Proceedings of the 25th IEEE EUROMICRO Conference* (pp. 326–334).
22. Kanasugi, A. (2003). A design of architecture for rough set processor. *Springer: Rough set theory and granular computing*.

23. Kanasugi, A., & Matsumoto, M. (2007). Design and implementation of rough rules generation from logical rules on FPGA board. *Springer: Rough sets and intelligent systems paradigms* (Vol. 4585, pp. 594–602). LNCS.
24. Sun, G., Qi, X., & Zhang, Y. (2011). A FPGA based implementation of rough set theory. In *Proceedings of Control and Decision Conference (CCDC)* (pp. 2561–2564).
25. Tiwari, K. S., & Kothari, A. (2015). Design and implementation of rough set co-processor on FPGA. *ICIC International*, 11(2).
26. Stepaniuk, J., Kopczynski, M., & Grzes, T. (2013). The first step toward processor for rough set methods. *Fundamenta Informaticae*, 127(1), 429–443.
27. Grze, T., Kopczynski, M., & Stepaniuk, J. (2013). FPGA in rough set based core and reduct computation. *Springer: Rough sets and knowledge technology* (pp. 263–270).
28. Kopczynski, M., Grzes, T., & Stepaniuk, J. (2014). Generating core in rough set theory: Design and implementation on FPGA. *Springer: Rough sets and intelligent systems paradigms* (pp. 209–216).
29. ThamaraiSelvi, S. (2010). Estimating job execution time and handling missing job requirements using rough set in grid scheduling. In *International Conference On Computer Design and Applications*, June 2010.
30. Advani, J. (2017). *Code profiling for RST algorithm on DSP and embedded processors*. M.Eng. thesis, E & TC Department College of Engineering Pune.

# A Comparison of the Effectiveness of Two Novel Clustering-Based Heuristics for the $p$ -Centre Problem



Mahima Yadav and V. Prem Prakash

**Abstract** Given a set of  $n$  demand points, the objective of the  $p$ -centre problem is to identify a subset of the demand points having  $p \ll n$  nodes (called centres) such that the maximum distance of any demand point to its nearest centre is minimized. The problem is NP-hard and finds application in facility location. This paper presents two novel heuristics for the  $p$ -centre problem that requires  $O(n^3)$  time. One of these is a deterministic heuristic that uses a minimum spanning tree-based clustering approach, and the other is a randomized heuristic that uses greedy clustering. Bounds on the computational time requirements of both heuristics are proved. The relative performance of the two heuristics is evaluated in the course of several computational experiments on a wide range of benchmark problems used in the literature for the  $p$ -centre problem.

**Keywords**  $p$ -centre · Heuristic · NP-hard · Facility location · Randomized

## 1 Introduction

Consider an undirected, weighted graph  $G = (V, E)$ , where the set  $V$  represents the vertex set of  $n = |V|$  vertices, and the set  $E$  represents the set of edges between these vertices. Also, let the length of the shortest path from vertex  $v_i$  to vertex  $v_j$  be denoted by  $d(v_i, v_j)$ . The goal of the  $p$ -centre problem is to find a subset  $W$  of  $V$  that contains exactly  $p$  nodes, or vertices, such that the largest distance from a vertex  $x \in \{V - W\}$  to its closest vertex  $w \in W$  is minimized, that is, to find the subset  $W = \{w_1, w_2, \dots, w_p\}$  of  $V$  that minimizes  $\max_{i \in \{1, \dots, n\}} \{\min_{j \in \{1, \dots, p\}} d(v_i, v_j)\}$ .

The  $p$ -centre problem is NP-hard [1]. An application of the problem arises in facility location, in which the goal is to select  $p$  facilities on a network in such a

---

M. Yadav (✉) · V. Prem Prakash  
Department of Electrical Engineering, Faculty of Engineering, Dayalbagh Educational Institute  
(Deemed University), Dayalbagh, Agra, India  
e-mail: [mahima.yadav72@gmail.com](mailto:mahima.yadav72@gmail.com)

V. Prem Prakash  
e-mail: [vpremprakash@dei.ac.in](mailto:vpremprakash@dei.ac.in)

manner that the maximum distance between clients and their corresponding nearest facilities are minimized. The problem also finds application in wireless communication systems, wherein a set of  $p$ -base stations need to be placed so as to provide maximum possible coverage. Several constrained real-world problems that require locating facilities like police- and fire-stations also map into the  $p$ -centre problem [2].

This work presents two novel fast polynomial-time heuristics for the  $p$ -centre problem with provable bounds on computational time. The first of these is a deterministic heuristic that uses a Minimum Spanning Tree (MST)-based clustering approach, whereas the second is a randomized heuristic. The relative performance of the two heuristics is studied in the course of several experiments on several standard benchmark problems used widely in the literature.

The rest of the paper is organized as follows. Section 2 gives a brief overview of earlier work in the literature for the  $p$ -centre problem. Section 3 describes the proposed novel heuristics, and provides some relevant theoretical results. Computational results are presented and discussed in Sect. 4. Concluding remarks are made in Sect. 5.

## 2 Related Work

Some exact approaches are given in the literature for special cases of the  $p$ -centre problem, c.f. the works of [3–5]. Several metaheuristic strategies are also proposed for solving the  $p$ -centre problem. Caruso et al. gave an algorithm [6] that solves a succession of set-covering problems using some preset parameters to find a solution to the  $p$ -centre problem. Mladenovic et al. [7] presented a couple of tabu search-based algorithms and a variable neighbourhood search-based approach. A scatter search-based approach is given in [8]. Chen and Chen [9] gave a novel relaxation algorithm. A bee colony optimisation (BCO) approach, called BCOi (BCO with improvement) is given by Davidovic et al. [10]. Jayalakshmi and Singh [2] proposed an artificial bee colony (ABC) algorithm and an invasive weed optimisation (IWO) algorithm for the problem.

## 3 Proposed Novel Heuristics

The two novel heuristics proposed in this work are described in greater detail in this section. The first heuristic presented here is a deterministic heuristic that uses an MST-based clustering approach to create  $p$ -clusters of demand points. From each cluster in turn, the heuristic then identifies one of the  $p$ -centres, chosen as the demand point that minimizes the maximum cost/distance from all other demand points in that cluster. The second heuristic takes a randomized approach to the selection of the  $p$ -centres. It then uses a greedy approach to create  $p$ -clusters and further improve the

choice of  $p$ -centres. In order to reduce the probability of poor random choices of  $p$ -centres, the heuristic repeats these steps  $n$  times and returns the best result obtained. Similar clustering based and randomized approaches have been successfully applied in heuristic design for other problems in the literature (c.f. [11–15]).

### 3.1 $p$ -Centres-Based Minimum Spanning Tree Clustering (pCMSTC) Heuristic

The  $p$ -centres-based Minimum Spanning Tree Clustering (pCMSTC) heuristic is a deterministic heuristic that solves the  $p$ -centre problem in  $O(n^3)$  time. The heuristic begins by constructing a matrix,  $D$  of the shortest distances between each pair of graph vertices. This is accomplished using the well-known All Pairs Shortest Paths (Floyd-Warshall) algorithm [16]. Thereafter, the heuristic iterates a fixed number of times. The number of iterations is determined by two empirically chosen parameters, the *threshold\_value* and *step\_size*. The *threshold\_value* represents the minimum number of nodes that a cluster should have and is varied from 1 to  $n/p$ , where  $n = |V|$  is the number of nodes, and  $p$  the number of central vertices, or centres. The *step\_size* is a small integer constant in the range [1, 30] that is experimentally determined for different instance sizes. In each iteration, a minimum spanning tree is constructed from  $D$  using Kruskal's algorithm [17]. Then,  $p-1$  edges are removed in decreasing order of edge weight from the MST, so as to obtain  $p$ -clusters. Care is taken to ensure that the number of nodes in each cluster has greater than or equal to *threshold\_value* nodes. The centre of each cluster is then computed by iteratively setting each cluster vertex in turn and setting as cluster centre, the vertex for which the highest edge weight to any other vertex of that cluster is minimal. Once all the  $p$ -cluster centres are identified, the value of  $\min p$  is computed. If this value is lower (better) than the global best value ( $\text{bestminp}$ ), then it replaces the  $\text{bestminp}$ . This entire process is iterated in a fixed number of times (as determined by the *threshold\_value* and *step\_size* parameters) and the best value of objective function ( $\text{bestminp}$ ) returned by the heuristic. Pseudocode for the heuristic is given in Fig. 1, and the computational time requirements are proved in Lemma 1.

**Lemma 1** The pCMSTC heuristic requires  $O(n^3)$  computation time on an input graph  $G = (V, E)$  having  $n = |V|$  demand points.

*Proof* The time required for constructing the matrix  $D$  is  $O(n^3)$ , since this step of the heuristic (step 1) is implemented using the All Pairs Shortest Paths (Floyd-Warshall) algorithm. The computation cost of constructing an MST (step 2) using Kruskal's algorithm is  $O(m \log n)$ , for  $m$  edges and  $n$  demand points. The inner loop (steps 5–12) iterates on the edges of the MST and removes  $p-1$  edges in descending order of weight. After removing each edge, a check is made to ensure that the number of nodes in the resultant sub-trees satisfies the threshold limit. This step is performed using a simple depth first traversal of the sub-trees, and has linear complexity. For a

**Fig. 1** Pseudocode for the pCMSTC heuristic

**Algorithm : pCMSTC(p)**

Let  $e$  represent an edge in the constructed **MST**  
 $p$  represent the number of centres  
 $n$  represent the number of demand points (nodes)  
 $threshold\_value$  represent the threshold value i.e.,  
number of nodes in a cluster  
 $step\_size$  represent the step size  
 $n_c$  be the number of clusters

1. Construct all pairs shortest distance matrix **D**
2. Build **MST** from **D**
3.  $threshold\_value = 1$ ;  $n_c = 0$
4. **repeat**
5.     **repeat**
6.         **for** each  $e \in \mathbf{MST}$  **do**
7.             Remove  $e$  from **MST** or (current longest edge)
8.             **if** (nodes in subtree  $< threshold\_value$ )
9.                 Restore  $e$  in **MST**
10.         **else**
11.              $n_c = n_c + 1$
12.         **until**  $n_c = p - 1$
13.     **for** each cluster  $Cluster_i$  **do** ( $0 \leq i \leq p - 1$ )
14.         Set as cluster center, the vertex  $u \in Cluster_i$  s.t.  
max.  $\{dist(u,v)\}$ ,  $v \in Cluster_i - \{u\}$  is minimized.
15.      $minp \leftarrow$  value of objective function
16.     **if**  $minp < bestminp$  **then**  
 $bestminp \leftarrow minp$
17.      $threshold\_value = threshold\_value + step\_size$
18. **until**  $threshold\_value = n/p$
19. **Return**  $bestminp$

total of  $n-1$  edges, the running time of the inner loop works out to  $O(n^2)$ . Updating cluster centres (steps 13, 14) are performed by repeated depth first traversals within each cluster and takes no more than  $O(n^2)$  time overall. The outer loop (steps 4–18) is run a constant number of times, as determined by the choice of the step size. Hence the overall computation time taken by the algorithm is  $O(n^3)$ .  $\square$

### 3.2 Randomized Greedy p-Centres (RGpC) Heuristic

The *Randomized Greedy p-centres (RGpC)* heuristic takes a randomized approach to solve the  $p$ -centre problem. The heuristic starts by randomly choosing  $p$ -centres from the vertex set  $V$ , and initializing each as the central vertex of its own cluster. Each of the remaining vertices is assigned to the cluster whose central vertex is reachable via the lowest cost path. Thereafter, within each cluster, each vertex of the cluster is set as the cluster centre in turn, and the vertex for which the maximum cost to any other node in the cluster is minimal is confirmed as the cluster centre. The objective function is then computed, and replaces the global best result if it is lower/better.

**Algorithm: RGpC(p)**

Let  $e$  represent an edge in the tree  
 $p$  represent the number of centres  
 $n$  represent the number of demand points (nodes)  
 $n_c$  be the number of clusters

1.  $i = 0$
2. **repeat**
3. Randomly select a set  $P$ , ( $|P|=p$ ), from vertex set  $V$
4. Create  $p$  clusters, with the central vertex of each initialized to a different vertex belonging to the set  $P$  of centers
5. **for** each vertex  $u \in V - P$  **do**
6. Assign  $u$  to the cluster for which  $\text{cost}(u, c)$  is minimal, where  $c$  is the corresponding cluster center
7. **for** each cluster **Cluster**, **do** ( $0 \leq i \leq p - 1$ )
8. Set as cluster center, the vertex  $u \in \text{Cluster}_i$  s.t.  $\max. \{\text{dist}(u, v)\}, v \in \text{Cluster}_i - \{u\}$  is minimized.
9.  $\text{minp} \leftarrow$  value of objective function
10. **if**  $\text{minp} < \text{bestminp}$  **then**  
 $\text{bestminp} \leftarrow \text{minp}$
11.  $i = i + 1$
12. **until**  $i = n$
13. **return**  $\text{bestminp}$

**Fig. 2** Pseudo-code for the *RGpC* heuristic

This procedure is repeated  $n = |V|$  times, and the global best solution obtained is returned by the heuristic. This heuristic also takes  $O(n^3)$  computation time. Pseudocode is given in Fig. 2, and bounds on computation time proved in Lemma 2.

**Lemma 2** The *RGpC* heuristic requires  $O(n^3)$  computation time, where  $n = |V|$  is the number of demand points represented by the input graph  $G = (V, E)$ .

*Proof* Selecting  $p$  random vertices as cluster centres and assigning the remaining  $|V| - p$  vertices to their respective clusters (steps 3–6 of pseudocode) require time that is asymptotically linear in  $n$ . Updating the cluster centres (steps 7–8) is performed by repeated depth first traversals within each cluster, and takes no more than  $O(n^2)$  time overall. Hence the time required for the inner loop is  $O(n^2)$ . The outer loop (steps 2–12) runs a total of  $n$  times, thus bringing the total computation time of the heuristic to  $O(n^3)$ .  $\square$

## 4 Experimental Work

The proposed heuristics are tested on several standard test problems used widely in the literature for the  $p$ -centre problem. These are listed in the Beasley OR Library [18] as instances of the uncapacitated  $p$ -median problem and comprise 40 sparse graphs with random edge weights. Each instance has a fixed number of demand



**Table 1** Results obtained on 40 benchmark instances having up to 900 demand points

S. No.	Inst. size ( $V$ )	No. of centres ( $p$ )	Objective function value	
			pCMSTC heuristic	RGpC heuristic
1	100	5	144	<b>127</b>
2	100	10	128	<b>112</b>
3	100	10	120	<b>109</b>
4	100	20	103	<b>92</b>
5	100	33	85	<b>73</b>
6	200	5	104	<b>92</b>
7	200	10	80	<b>77</b>
8	200	20	75	<b>73</b>
9	200	40	71	<b>60</b>
10	200	67	<b>42</b>	70
11	300	5	64	<b>62</b>
12	300	10	72	<b>72</b>
13	300	30	55	<b>49</b>
14	300	60	50	<b>47</b>
15	300	100	<b>37</b>	44
16	400	5	52	<b>49</b>
17	400	10	48	<b>47</b>
18	400	40	50	<b>50</b>
19	400	80	38	<b>35</b>
20	400	133	<b>30</b>	35
21	500	5	45	<b>42</b>
22	500	10	53	<b>47</b>
23	500	50	34	<b>33</b>
24	500	100	<b>33</b>	39
25	500	167	<b>24</b>	44
26	600	5	44	<b>41</b>
27	600	10	42	<b>39</b>
28	600	60	<b>36</b>	57
29	600	120	<b>24</b>	36
30	600	200	<b>21</b>	40
31	700	5	34	<b>31</b>
32	700	10	<b>40</b>	72
33	700	70	28	<b>24</b>
34	700	140	<b>23</b>	41
35	800	5	36	<b>32</b>

(continued)

**Table 1** (continued)

S. No.	Inst. size ( $V$ )	No. of centres ( $p$ )	Objective function value	
			pCMSTC heuristic	RGpC heuristic
36	800	10	<b>38</b>	42
37	800	80	30	<b>26</b>
38	900	5	<b>31</b>	40
39	900	10	<b>32</b>	74
40	900	90	26	<b>23</b>

points associated with it, and this number varies between 100 and 900. Specifically, there are five instances each of 100, 200, 300, 400, 500 and 600 demand points; four instances each having 700 demand points, and three instances having 800 and 900 demand points. The number of centres varies between 5 and 200. Both heuristics were implemented in C and tested on a computing system with an Intel core i5 processor and 4 GB of RAM.

The results of the computational work are presented in Table 1. In each row of the table, columns 1–3 represent the instance number, the number of demand points and the number of facility centres ( $p$ ), respectively. The highest cost from any demand point to its nearest facility is given under the objective function value column (no. 4) for the pCMSTC and RGpC heuristics.

On the smaller problem sizes, the RGpC heuristic obtains better (lower) results, vis-à-vis, the pCMSTC heuristic for almost all values of  $p$  (number of centres) considered. The pCMSTC heuristic performs better only when the number of demand points is very large. For instance, on 100 node instances, the RGpC heuristic obtains better results in all cases. On the 200, 300, 400 and 500 node instances also, it consistently performs better for almost all the instances, losing out to the pCMSTC heuristic only in the last case for each size of node instance, when  $p$  is very large. However, as the problem size grows larger, the heuristics obtain competitive results. The pCMSTC heuristic obtains better results on three of the 600 node instances, and both heuristics obtain better results on two instances each of the 700 and 800 node instances. On the 900 node instances, however, the pCMSTC heuristic obtains better performance than the RGpC heuristic. Overall, the randomized RGpC heuristic obtains better results on 27 instances, whereas the deterministic pCMSTC heuristic does better on 13 instances.

## 5 Conclusions

The  $p$ -centre problem is NP-hard and arises in problem domains such as wireless communications and facility location. This work evaluates the performance of two novel heuristics that take different approaches towards solving the  $p$ -centre problem. The first of the proposed heuristics, called pCMSTC, uses a deterministic, MST

clustering-based approach, while the second, called the RGpC heuristic, takes a randomized approach. The performance of the heuristics is compared to several benchmark problems. The RGpC heuristic clearly performs much better on the small-to-medium sized test instances, particularly when the number of centres ( $p$ ) is not high in comparison to the number of demand points. For such instances, the pCMSTC heuristic is seen to be more suitable when  $p$  is large. On large problem sizes, both heuristics are seen to be competitive, with the pCMSTC heuristic possibly obtaining slightly better performance as the problem size grows quite large.

## References

1. Kariv, O., & Hakimi, S. L. (1969). An algorithmic approach to network location problems. II: The  $p$ -medians. *SIAM Journal of Applied Mathematics*, 37, 539–560.
2. Jayalakshmi, B., & Singh, A. (2018). Two swarm intelligence-based approaches for the  $p$ -centre problem. *International Journal of Swarm Intelligence*, 3(4), 290–308.
3. Drezner, Z. (1984). The planar two center and two median problems. *Transportation Science*, 18, 451–461.
4. Handler, G. Y. (1990).  $p$ -center problems. In *Discrete location theory* (pp. 305–315). Wiley.
5. Daskin, M. S. (1995). *Network and discrete location: Models, algorithms, and application*. Wiley.
6. Caruso, C., Colomi, A., & Aloï, L. (2003). Dominant, an algorithm for the  $p$ -center problem. *European Journal of Operational Research*, 149(1), 53–64.
7. Mladenovic, N., Labbe, M., & Hansen, P. (2003). Solving the  $p$ -center problem with tabu search and variable neighborhood search. *Networks*, 42(1), 48–64.
8. Pacheco, J. A., & Casado, S. (2005). Solving two location models with few facilities by using a hybrid heuristic: A real health resources case. *Computers & Operations Research*, 32(12), 3075–3091.
9. Chen, D., & Chen, R. (2009). New relaxation-based algorithms for the optimal solution of the continuous and discrete  $p$ -center problems. *Computers & Operations Research*, 36(5), 1646–1655.
10. Davidovic, T., Ramljak, D., Selmic, M., & Teodorovic, D. (2011). Bee colony optimization for the  $p$ -center problem. *Computers & Operations Research*, 38(10), 1367–1376.
11. Julstrom, B. A. (2009). Greedy heuristics for the bounded diameter minimum spanning tree problem. *ACM Journal of Experimental Algorithmics*, 14(1), 1–14.
12. Patvardhan, C., & Prakash, V. P. (2009, December). Novel deterministic heuristics for building minimum spanning trees with constrained diameter. *Pattern Recognition and Machine Intelligence*, LNCS 5909 (pp. 68–73).
13. Patvardhan, C., Prakash, V. P., & Srivastav, A. (2014). Parallel heuristics for the bounded diameter minimum spanning tree problem. In *India Conference (INDICON), 2014 Annual IEEE*, 11–13 December 2014 (pp. 1–5). IEEE Press.
14. Patvardhan, C., Prakash, V. P., & Srivastav, A. (2015). Fast heuristics for large instances of the euclidean bounded diameter minimum spanning tree problem. *Informatica (Slovenia)*, 39(3), 281–292.
15. Prakash, V. P., Patvardhan, C., & Srivastav, A. (2018). Effective heuristics for the bi-objective euclidean bounded diameter minimum spanning tree problem. In P. Bhattacharyya, H. Sastry, V. Marriboyina, & R. Sharma (Eds.), *Smart and innovative trends in next generation computing technologies* (NGCT 2017). Singapore.
16. Floyd, R. W. (1962). Algorithm 97: Shortest path. *Communications of the ACM*, 5(6), 345.

17. Kruskal, J. B. (1956). On the shortest spanning subtree and the travelling salesman problem. In *Proceedings of the American Mathematical Society* (pp. 48–50).
18. Beasley, J. E. (1990). OR-library: Distributing test problems by electronic mail. *Journal of the Operational Research Society*, 41(11), 1069–1072.

# Half-Life Teaching Factor Based TLBO Algorithm



Ruchi Mishra, Nirmala Sharma and Harish Sharma

**Abstract** Teaching-learning-based optimization algorithm (TLBOA) is a significant metaheuristic algorithm. It is a proficient approach for solving multidimensional, linear, and nonlinear optimization problems. It is based on teaching-learning (TL) process that searches for a global optimum through two modules of learning: (a) teacher-phase (TP) and (b) learner-phase (LP). For avoiding the premature convergence of TLBOA, half-life teaching factor is discovered in this paper. The proposed strategy is known as half-life teaching factor based TLBO (HRTLBO) algorithm. The performance of HRTLBO is calculated over 20 benchmark functions and compared with various state-of-art algorithms namely, TLBOA, global-Best inspired biogeography-based optimization (GBBO), particle swarm optimization (PSO), and covariance matrix adaptation evolution strategy (CMA-ES). The obtained outcomes validate the authenticity of the discovered HRTLBO.

**Keywords** Teaching-learning-based optimization · Swarm intelligence based algorithm · Optimization

## 1 Introduction

Teaching-learning-based optimization algorithm (TLBOA) is basically a nature-inspired algorithm (NIA). Nature has been a source of inspiration for technological development. This results in the development of NIAs. It is a proficient approach for solving complex real-world optimization problems. Mainly NIAs is split into two types: swarm intelligence (SI) based optimization algorithms and evolutionary

---

R. Mishra (✉) · N. Sharma · H. Sharma  
Rajasthan Technical University, Kota, India  
e-mail: [ruchimishra428@gmail.com](mailto:ruchimishra428@gmail.com)

N. Sharma  
e-mail: [nsharma@rtu.ac.in](mailto:nsharma@rtu.ac.in)

H. Sharma  
e-mail: [hsharma@rtu.ac.in](mailto:hsharma@rtu.ac.in)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_25](https://doi.org/10.1007/978-981-15-0694-9_25)

algorithms. SI-based optimization algorithms are those which are inspired by the collective behavior of social insect colonies and other animal's societies. TLBOA is invented by Dr. Rao et al. [1] in the year 2011. It is a teaching-learning (TL) procedure inspired algorithm based on the effect of the sway of a teacher (supervisor) on the outturn of learners (scholars) in the class. The grades of the scholars which totally depends on the standard or caliber of a supervisor are considered in terms of an outturn of this algorithm. The supervisor is the person who trains scholars and helps them to acquire knowledge so that scholars score good marks and grades. Interaction with the other scholars is also helpful for improving their marks and grades. TLBOA describes through two modules of learning: (a) teacher-phase (TP) (where supervisor teaches scholars) and (b) learner-phase (LP) (where learning is through the synergistic good communication among the scholars). Researchers have been working to enhance the functioning of TLBOA [2–4]. It is reported in the literature that TLBOA has premature convergence problem [5].

In the above context, this article proposes a modified variant of TLBOA. A half-life formula based teaching factor is introduced in the TLBOA. This factor reduces the step size of the solutions and helps in reducing the premature convergence. The algorithm is titled as half-life teaching factor based TLBO (HLTLBO) algorithm. This proposed variant is compared with TLBOA, global-best inspired biogeography-based optimization (GBBO) [6], particle swarm optimization (PSO) [7], and covariance matrix adaptation evolution strategy (CMA-ES) [8]. The obtained outcomes prove the authenticity of the discovered approach.

The other sections are organized as follows: In Sect. 2, TLBOA is discussed. In Sect. 3, we introduce our discovered HLTLBO algorithm and for measuring the performance of HLTLBO algorithm, a comparison has been made with various algorithms in Sect. 4. At last, Sect. 5 includes the conclusion of the proposed work.

## 2 Teaching-Learning-Based Optimization Algorithm

TLBOA illustrates the classroom department of supervisor and scholars. The functioning of TLBOA is alienated into two parts, TP and LP [9] which is elucidated beneath.

### 2.1 Teacher-Phase

During this phase the supervisor trains scholars so that scholars can achieve better outcomes regarding grades and marks in subject matters. The supervisor is a person who is highly educated, experienced, and knowledgeable. The very first step of TP is to initialize the total number of scholars (population) and total subjects (design variables). Let “ $M$ ” be the total subjects, “ $N$ ” be the number of scholars ( $k = 1, 2, 3, \dots, N$ ), and  $M_{ji}$  be the mean result of the scholars in a particular subject matters “ $j$ ” ( $j = 1, 2, 3,$

... , M) at any iteration (*it*)  $X_i$ . The supervisor will put maximum efforts to increase knowledge level of class, whereas scholars gain knowledge regarding the caliber of teaching delivered by a supervisor and the caliber of scholars present in the classroom. Considering this facet, the difference between the outcome of supervisor and scholar’s mean result in particular subject matters is elucidated by

$$Difference\_Mean_{ji} = R_i \times (M_T - T_F \times M_{ji}) \tag{1}$$

where  $M_T$  is defined as best scholar in subject matter  $j$ .  $R_i$  is the randomly selected no. in the Range [0, 1] and  $T_F$  is the Teaching factor which adjudicates the value of mean to be changed.  $T_F$  is not a parametric quantity of the TLBOA, it can be either 1 or 2 and it is adjudicated randomly with equal chances as

$$T_F = round[1 + R(0, 1)\{2 - 1\}] \tag{2}$$

The position update equation of the old solution in TP is articulated by

$$X_{newvalue} = X_{oldvalue} + Difference\_Mean_{ji} \tag{3}$$

where  $X_{newvalue}$  is the updated value of  $X_{oldvalue}$ .  $X_{newvalue}$  is acceptable, when its outcome is better than  $X_{oldvalue}$ . If the outcome is not better than the  $X_{oldvalue}$ ,  $X_{newvalue}$  is neglected. After that, selected values in TP are reckoned as *input* to the LP.

## 2.2 Learner-Phase

LP is the next phase of an algorithm where a particular scholar gain knowledge from other scholars too if they have some improvement ideas related to subject matters. Any two scholars  $X_{p1}$  and  $X_{q1}$  are randomly selected from the population “ $N$ ”, such that  $p1 \neq q1$ . The updated parameter  $X_{newvalue}$  is described by the Eq. 4.

$$\begin{aligned} & \text{if } f(X_{p1}) < f(X_{q1}) \\ & X_{newvalue} = X_i + R_i \times (X_{p1} - X_{q1}) \\ & \text{else} \\ & X_{newvalue} = X_i + R_i \times (X_{q1} - X_{p1}) \end{aligned} \tag{4}$$

The TLBOA is pictured through a flowchart given in Fig. 1, in which first initializes the total number of scholars and this flow go forward toward the TP and the outturn of TP is farther provided as *input* for LP that emits the most excellent value of the objective function.

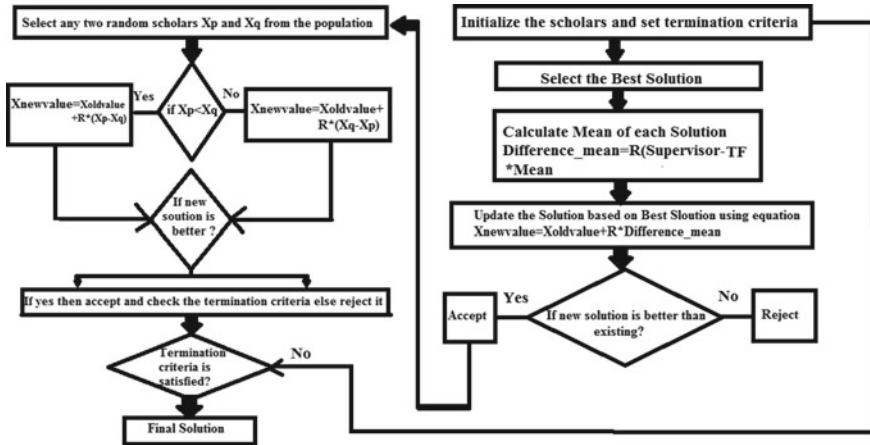


Fig. 1 TLBO algorithm

### 3 Half-Life Teaching Factor Based TLBO Algorithm

It is reported in the literature that TLBOA suffers from the problem of premature convergence [5]. To overcome the above problem, an updated  $T_F$  is proposed that is inspired by Half-life formula. The proposed algorithm is known as half-life teaching factor based TLBO (HLTLBO) algorithm.

**Half-life formula** is the time required for the amount of something to fall to half its initial value. The mathematical representation of half-life is given below

$$t_{\frac{1}{2}} = \frac{\ln 2}{\lambda} = \frac{0.693}{\lambda} \tag{5}$$

where,  $t_{\frac{1}{2}}$  is half-life and  $\lambda$  is disintegration constant. This formula is used in  $T_F$  to reduce the distance to its half, between the position of the supervisor and the scholars or more specifically, it can be described as an innovative approach to minimize the time required by the scholars to grasp the knowledge, provide by their supervisor. Based on the Eq. 5, the  $T_F$  is calculated as per the Eq. 2

$$T_F = \frac{0.693}{\lambda} = \frac{0.693}{[(MaxIt - it)/MaxIt] + 1} \tag{6}$$

where  $it$  is the current generation and  $MaxIt$  is the total count of generations. As per the Eq. 2, the value of  $T_F$  which is inspired from Eq. 5 increases in a fast rate during the early iterations and after a lapse of few iterations, it decreases in a slow rate that helps us in avoiding the premature convergence.

Figure 2 depicts the value of  $T_F$  with respect to iterations.



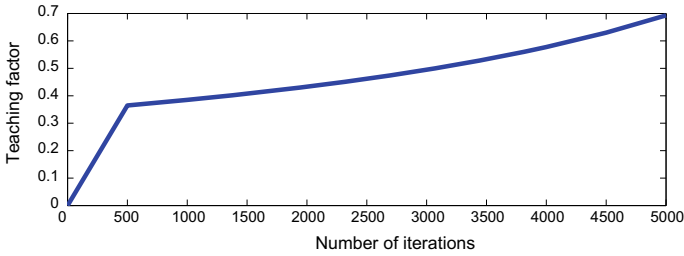


Fig. 2 Graph between teaching factor and various iterations

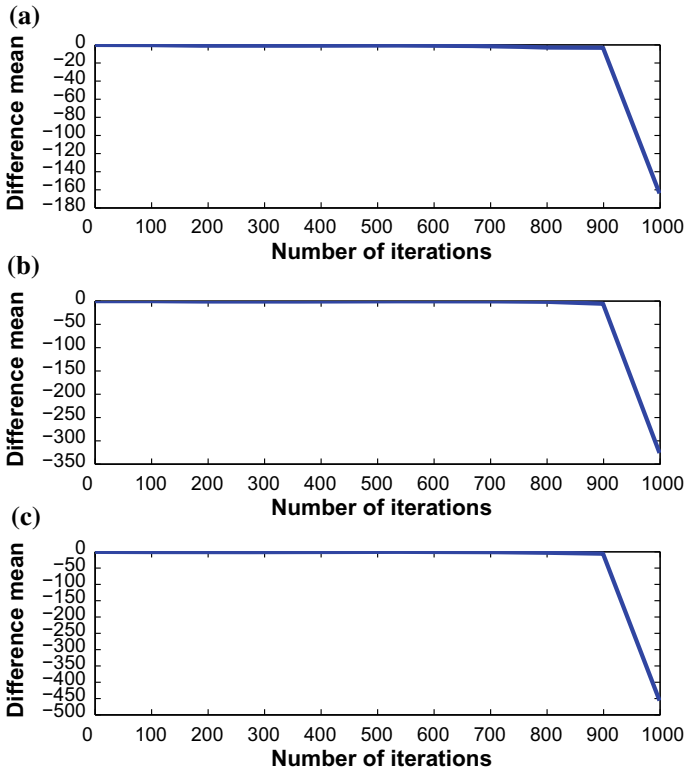


Fig. 3 DM for test functions; **a** DM for first dimension, **b** DM for second dimension, and **c** DM for third dimension

A three-dimensional problem of Parabola/Sphere is considered as an example for calculating the difference mean (DM) (DM represents the step size for the solutions) as per the Eq. 3. Figure 3a-c shows the graphs that represents the DM for all the three dimensions.

It is clear from Fig. 3 that during last iterations, the value of difference mean has been reduced so it is clear that step size will also reduce and convergence speed becomes fast. It avoids the premature convergence of the algorithm. The fundamental contribution of discovered variant avoids premature convergence of TLBOA.

## 4 Experimental Results

### 4.1 Considered Test Problems

To estimate the competence of an anticipated algorithm HLTLBO, 20 various global optimization problems ( $fun_1$  to  $fun_{20}$ ) are chosen from CEC2005 [10], CEC2008 [11], and CEC2013 [12]. These problems have different complexity levels. Test problems  $fun_1$  to  $fun_{20}$  are reckoned along with their offset values shown in Table 1.

### 4.2 Experimental Setting

To estimate the functioning of the propounded algorithm HLTLBO, a comparison is made among TLBOA [1], GBBO [6], PSO [7], and CMA-ES [8]. To test them over the considered problems, an experiment environment is created where

- Totalrun = 30,
- Population size  $N = 50$ ,
- $MaxIt = 5000$ ,
- $TF = \frac{0.693}{[(MaxIt-it)/MaxIt]+1}$ ,
- $R_i = R[0, 1]$ .

### 4.3 Results Comparison

Table 2 endows with a report of the comparison made between the well thought-out algorithms. Table 2 presents a record of the *standard\_deviation* (SD), *mean\_error* (ME), *success\_rate* (SR), and average count for function evaluations (AFEs). For reaching the termination criteria by the algorithm, *AFEs* is called in 100 runs and for achieving the optima by the algorithm, *acceptable\_error* (AE) is called in 30 runs. After calculating the outcomes, it can be accomplished that HLTLBO performs better than TLBOA [1], GBBO [6], PSO [7], and CMA-ES [8] regarding proficiency, reliability, and accuracy. Further, a nonparametric test named as *Mann-Whitney U rank sum test* [13] and various statistical tests like *AR-test* and *Boxplots* [14] are constructed to analyze the results more intensively.

**Table 1** Test problems: *TP*, *D*: dimension, *AE*: acceptable error

S.no	Test problem	Objective function	Search range	D	AE
1	De Jong f4	$fun_1(x) = \sum_{i=1}^D i \cdot (x_i)^4$	[-5.12, 5.12]	30	1.0E-05
2	Griewank	$fun_2(x) = 1 + \frac{1}{4000} \sum_{i=1}^D x_i^2 - \prod_{i=1}^D \cos(\frac{x_i}{\sqrt{i}})$	[-600, 600]	30	1.0E-05
3	Rosenbrock	$fun_3(x) = \sum_{i=1}^D (100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2)$	[-30, 30]	30	1.0E-02
4	Rastrigin	$fun_4(x) = 10D + \sum_{i=1}^D [x_i^2 - 10\cos(2\pi x_i)]$	[-5.12, 5.12]	30	1.0E-03
5	Ackley	$fun_5(x) = -20 + e + \exp(-\frac{0.2}{D} \sqrt{\sum_{i=1}^D x_i^3})$	[-1, 1]	30	1.0E-05
6	Alpine	$fun_6(x) = \sum_{i=1}^D ( x_i \sin x_i + 0.1x_i )$	[-10, 10]	30	1.0E-05
7	Cosine mixture	$fun_7(x) = \sum_{i=1}^D x_i^2 - 0.1(\sum_{i=1}^D \cos 5\pi x_i) + 0.1D$	[-1, 1]	30	1.0E-05
8	Exponential	$fun_8(x) = -(\exp(-0.5 \sum_{i=1}^D x_i^2)) + 1$	[-1, 1]	30	1.0E-05
9	Zakharov	$fun_9(x) = \sum_{i=1}^D x_i^2 + (\sum_{i=1}^D \frac{ix_i}{2})^2 + (\sum_{i=1}^D \frac{ix_i}{2})^4$	[-5.12, 5.12]	30	1.00E-02
10	Brown3	$fun_{10}(x) = \sum_{i=1}^{D-1} (x_i^{2(x_{i+1})^2+1} + x_{i+1}^{2x_i^2+1})$	[-1, 4]	30	1.0E-05
11	Schewel prob 3	$fun_{11}(x) = \sum_{i=1}^D  x_i  + \prod_{i=1}^D  x_i $	[-10, 10]	30	1.0E-05
12	Axis parallel hyper-ellipsoid	$fun_{12}(x) = \sum_{i=1}^D ix_i^2$	[-5.12, 5.12]	30	1.0E-05
13	Sum of different powers	$fun_{13}(x) = \sum_{i=1}^D  x_i ^{i+1}$	[-1, 1]	30	1.0E-05
14	Step function	$fun_{14}(x) = \sum_{i=1}^D ([x_i + 0.5])^2$	[-100, 100]	30	1.00E-05
15	Rotated hyper-ellipsoid	$fun_{15}(x) = \sum_{i=1}^D \sum_{j=1}^i x_j^2$	[-65.536, 65.536]	30	1.0E-05
16	Shifted sphere	$fun_{16}(x) = \sum_{i=1}^D z_i^2 + f_{bias}$ , $z = x - o, x = [x_1, x_2, \dots, x_D]$ , $o = [o_1, o_2, \dots, o_D]$	[-100, 100]	100	1.0E-01
17	Shifted Griewank	$fun_{17}(x) = \sum_{i=1}^D \frac{z_i^2}{4000} - \prod_{i=1}^D \cos(\frac{z_i}{\sqrt{i}}) + 1 + f_{bias}$ , $z = (x - o), x = [x_1, x_2, \dots, x_D]$ , $o = [o_1, o_2, \dots, o_D]$	[-600, 600]	100	1.0E-01
18	Dekkers and Aarts	$fun_{18}(x) = 10^5 x_1^2 + x_2^2 - (x_1^2 + x_2^2)^2 + 10^{-5} (x_1^2 + x_2^2)^4$	[-20, 20]	2	5.0E-01
19	Meyer and Roth	$fun_{19}(x) = \sum_{i=1}^5 \left( \frac{x_1 x_3 x_i}{1+x_1 t_i + x_2 v_i} - y_i \right)^2$	[-10, 10]	3	1.95E-02
20	Pressure vessel	$fun_{20}(x) = 0.6224x_1 x_3 x_4 + 1.7781x_2 x_3^2 + 3.1661x_1^2 x_4 + 19.84x_1^2 x_3$	$x_1 = [1.1, 12.5]$ $x_2 = [0.6, 12.5]$ $x_3 = [0, 240]$ $x_4 = [0, 240]$	4	1.0E-05

**Table 2** Comparison of the results of test functions, *TP*: test problem

TP	Algorithm	SD	ME	AFE	SR
<i>fun</i> <sub>1</sub>	HLTLBO	2.13E−06	6.16E-06	2696.67	30.00
	TLBOA	2.18E−06	6.41E−06	2803.33	30.00
	GBBO	1.16E−06	8.86E−06	23302.00	30.00
	PSO	1.25E−06	8.49E−06	6450.00	30.00
	CMA-ES	1.96E−06	7.37E−06	20908.00	30.00
<i>fun</i> <sub>2</sub>	HLTLBO	1.18E−06	8.14E−06	4210.00	30.00
	TLBOA	1.35E−06	8.06E−06	4403.33	30.00
	GBBO	8.46E−04	7.61E−01	200000.00	0.00
	PSO	2.81E−03	7.60E−01	200000.00	0.00
	CMA-ES	0.00E+00	7.59E−01	200128.00	0.00
<i>fun</i> <sub>3</sub>	HLTLBO	4.21E−01	2.45E+01	200050.00	0.00
	TLBOA	1.32E+00	1.57E+01	200050.00	0.00
	GBBO	4.12E+02	1.88E+02	200000.00	0.00
	PSO	1.48E+01	2.18E+01	200000.00	0.00
	CMA-ES	7.53E+02	3.09E+02	200128.00	0.00
<i>fun</i> <sub>4</sub>	HLTLBO	5.38E+00	7.05E+00	179663.33	8.00
	TLBOA	5.11E+00	1.06E+01	197860.00	1.00
	GBBO	1.02E+01	3.90E+01	200000.00	0.00
	PSO	1.49E+01	5.28E+01	200000.00	0.00
	CMA-ES	9.05E+00	1.54E+02	200128.00	0.00
<i>fun</i> <sub>5</sub>	HLTLBO	6.69E−07	8.49E−06	7203.33	30.00
	TLBOA	6.46E−07	8.77E−06	7570.00	30.00
	GBBO	5.20E−03	2.67E−02	200000.00	0.00
	PSO	8.26E−01	7.79E−01	107158.33	15.00
	CMA-ES	3.49E−07	9.46E−06	64150.33	30.00
<i>fun</i> <sub>6</sub>	HLTLBO	7.68E−07	8.91E−06	7106.67	30.00
	TLBOA	6.00E−07	8.74E−06	7406.67	28.00
	GBBO	2.90E−03	8.40E−03	200000.00	0.00
	PSO	5.51E−04	1.40E−04	86393.33	19.00
	CMA-ES	6.26E−07	9.26E−06	71597.00	30.00
<i>fun</i> <sub>7</sub>	HLTLBO	1.12E−06	7.73E−06	3843.33	30.00
	TLBOA	1.51E−06	7.81E−06	4030.00	29.00
	GBBO	2.93E−01	8.77E−01	200000.00	0.00
	PSO	4.28E−01	9.95E−01	200000.00	0.00
	CMA-ES	7.61E−07	8.98E−06	32296.00	27.00
<i>fun</i> <sub>8</sub>	HLTLBO	1.33E−06	7.55E-06	2983.33	30.00
	TLBOA	1.15E−06	7.72E−06	3103.33	30.00
	GBBO	6.25E−07	9.50E−06	6537.67	30.00
	PSO	2.50E−06	7.47E−06	7320.00	30.00
	CMA-ES	1.18E−06	8.61E−06	22933.33	30.00

(continued)

**Table 2** (continued)

TP	Algorithm	SD	ME	AFE	SR
<i>fun<sub>9</sub></i>	HLTLBO	1.10E-03	8.70E-03	23110.00	30.00
	TLBOA	7.35E-04	9.10E-03	23990.00	29.00
	GBBO	1.10E-03	1.00E-02	137120.00	27.00
	PSO	3.63E-04	9.50E-03	52337.00	30.00
	CMA-ES	4.56E+01	1.76E+02	200128.00	0.00
<i>fun<sub>10</sub></i>	HLTLBO	9.90E-07	7.21E-06	3736.67	30.00
	TLBOA	1.45E-06	7.41E-06	3960.00	30.00
	GBBO	5.02E-06	1.78E-05	199925.00	1.00
	PSO	1.75E-06	8.16E-06	7933.33	30.00
	CMA-ES	1.25E-06	8.40E-06	32221.33	30.00
<i>fun<sub>11</sub></i>	HLTLBO	7.19E-07	8.77E-06	7650.00	30.00
	TLBOA	6.73E-07	8.97E-06	8036.67	30.00
	GBBO	6.80E-03	3.32E-02	200000.00	0.00
	PSO	4.47E-02	1.32E-02	158420.00	7.00
	CMA-ES	5.08E-07	9.40E-06	73331.67	28.00
<i>fun<sub>12</sub></i>	HLTLBO	1.41E-06	8.35E-06	4430.00	30.00
	TLBOA	1.46E-06	7.04E-06	4653.33	30.00
	GBBO	2.41E-04	5.22E-04	200000.00	0.00
	PSO	1.45E-06	8.28E-06	38152.33	30.00
	CMA-ES	8.90E-07	8.89E-06	38152.67	25.00
<i>fun<sub>13</sub></i>	HLTLBO	2.67E-06	5.37E-06	1916.67	30.00
	TLBOA	2.55E-06	6.38E-06	1920.00	30.00
	GBBO	2.26E-06	7.49E-06	4306.33	30.00
	PSO	2.04E-06	7.49E-06	5243.67	30.00
	CMA-ES	3.27E-06	7.05E-06	51777.33	25.00
<i>fun<sub>14</sub></i>	HLTLBO	0.00E+00	0.00E+00	2286.67	30.00
	TLBOA	0.00E+00	0.00E+00	2360.00	30.00
	GBBO	0.00E+00	0.00E+00	52217.00	0.00
	PSO	1.92E-06	0.00E+00	15567.33	30.00
	CMA-ES	0.00E+00	0.00E+00	16377.67	30.00
<i>fun<sub>15</sub></i>	HLTLBO	1.24E-06	7.88E-06	5023.33	30.00
	TLBOA	1.24E-06	8.27E-06	5236.67	30.00
	GBBO	2.50E-03	5.40E-03	200000.00	30.00
	PSO	1.09E-11	7.56E-06	7300.00	30.00
	CMA-ES	9.50E-07	8.56E-06	42998.33	30.00
<i>fun<sub>16</sub></i>	HLTLBO	5.49E-07	9.21E-06	25896.67	30.00
	TLBOA	1.26E-06	8.80E-06	200050.00	0.00
	GBBO	2.34E-06	7.96E-06	190210.33	4.00
	PSO	1.75E+02	1.75E+04	200000.00	0.00
	CMA-ES	1.91E-06	7.26E-06	96653.67	30.00

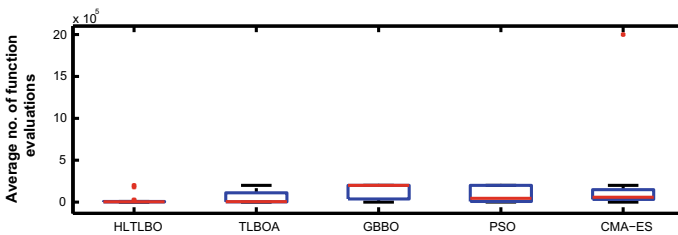
(continued)

**Table 2** (continued)

TP	Algorithm	SD	ME	AFE	SR
<i>fun</i> <sub>17</sub>	HLTLBO	8.12E+01	6.67E+02	150.00	30.00
	TLBOA	1.06E+02	6.28E+02	200050.00	0.00
	GBBO	9.85E+01	6.16E+02	171.33	30.00
	PSO	9.83E−06	3.62E+02	310.33	30.00
	CMA-ES	1.61E+02	4.12E+02	171.33	30.00
<i>fun</i> <sub>18</sub>	HLTLBO	6.96E+03	9.94E+03	376.67	30.00
	TLBOA	5.95E+03	8.34E+03	383.33	30.00
	GBBO	6.04E+03	3.91E+03	575.00	30.00
	PSO	1.17E+03	1.32E+03	910.00	30.00
	CMA-ES	7.10E+03	1.06E+04	35219.67	30.00
<i>fun</i> <sub>19</sub>	HLTLBO	2.11E−03	1.45E−02	153.33	30.00
	TLBOA	1.41E−03	1.37E−02	200050.00	0.00
	GBBO	5.21E−18	1.26E−02	200000.00	0.00
	PSO	5.20E−18	1.26E−02	200000.00	0.00
	CMA-ES	2.59E−06	3.56E−06	2000128.00	0.00
<i>fun</i> <sub>20</sub>	HLTLBO	4.05E+38	1.94E+39	150.00	30.00
	TLBOA	5.49E+38	1.84E+39	150.00	30.00
	GBBO	6.50E−04	2.00E−03	200000.00	0.00
	PSO	1.69E−16	8.34E−16	20020.00	30.00
	CMA-ES	1.20E−16	8.36E−16	91779.67	25.00

### 4.4 Statistical Analysis

The *Boxplots* are generated for the average count for function evaluations (AFEs) for performing the overall comparison. Basically, *Boxplots* are used for showing the empirical distribution of the data effectively. The *Boxplots* for HLTLBO, TLBOA [1], GBBO [6], PSO [7], and CMA-ES [8] are presented in Fig. 4 which clearly pictures that HLTLBO is very effectual regarding a function evaluations because interquartile



**Fig. 4** *Boxplots* graph for AFEs

range is very low and also a median is very low for HLTLBO as compared to TLBOA, GBBO, PSO, and CMA-ES.

Further, measuring the convergence speed regarding average count for function evaluations (AFEs), acceleration test (AR) has been executed. AR is calculated as

$$AR = \frac{AFE_{ALGO}}{AFE_{HLTLBO}} \tag{7}$$

Here,  $ALGO \in \{TLBOA, GBBO, PSO, \text{ and } CMA - ES\}$  and the value of AR is greater than 1 which shows that HLTLBO is more faster than other algorithms. Table 3 shows a differentiation between the HLTLBO and TLBOA, HLTLBO and GBBO, HLTLBO and PSO, and HLTLBO and CMA-ES regarding AR.

A well-known nonparametric test named as *Mann–Whitney U rank sum test* is performed at a five percent significance level ( $\alpha = 0.05$ ) between HLTLBO—TLBOA, HLTLBO—GBBO, HLTLBO—PSO, and HLTLBO—CMA-ES presented in Table 4.

**Table 3** Acceleration rate (AR) of HLTLBO compared to the TLBOA, GBBO, CMA – ES, and PSO

Test problems	TLBOA	GBBO	PSO	CMA-ES
<i>fun</i> <sub>1</sub>	1.0395549811	8.64103821211	2.39184175045	7.75327555356
<i>fun</i> <sub>2</sub>	1.0459223998	47.5059382423	47.5059382423	47.5363420428
<i>fun</i> <sub>3</sub>	1.0000000000	0.99750062567	0.99750062567	1.00038990255
<i>fun</i> <sub>4</sub>	1.1012820334	1.11319319499	1.11319319499	1.11390563856
<i>fun</i> <sub>5</sub>	1.0509023667	27.7649237749	14.8762147177	8.90564557112
<i>fun</i> <sub>6</sub>	1.0422138835	28.1425889862	12.1566603553	10.0746247182
<i>fun</i> <sub>7</sub>	1.0485684547	52.0381617696	52.0381617696	8.40312236267
<i>fun</i> <sub>8</sub>	1.0422346411	2.19139668378	2.45363131233	7.92626736098
<i>fun</i> <sub>9</sub>	1.0380787538	5.93336218090	2.26469061011	8.65980095223
<i>fun</i> <sub>10</sub>	1.0597680548	53.5035677654	2.12310434322	8.62301507922
<i>fun</i> <sub>11</sub>	1.0505446667	26.1437908497	20.7084967322	9.58583878433
<i>fun</i> <sub>12</sub>	1.0504138375	45.1467262863	8.61226485334	8.61234011298
<i>fun</i> <sub>13</sub>	1.0017391133	2.13913039766	2.73582605689	27.0142603824
<i>fun</i> <sub>14</sub>	1.0320699558	22.8354224076	6.80787160634	7.16224480811
<i>fun</i> <sub>15</sub>	1.0424684944	39.8142006623	1.45321832423	8.55972135087
<i>fun</i> <sub>16</sub>	7.7249324138	7.34497360233	7.72300166344	3.73228214355
<i>fun</i> <sub>17</sub>	133333.66666	1.14222200000	2.06888866677	1.14222222223
<i>fun</i> <sub>18</sub>	1.01769893655	1.52654853755	2.41592898977	93.5035316369
<i>fun</i> <sub>19</sub>	1304.67419678	1304.34810965	1304.34810964	13044.3158792
<i>fun</i> <sub>20</sub>	1.0000000000	1333.33333334	133.466666667	611.864444667

**Table 4** Comparison based on *AFE* and the *Mann–Whitney U rank sum test* at (“+ve” indicates HLTLBO is better, “–ve” indicates HLTLBO is worse and “=” indicates that there is no noticeable difference)

Function no	HLTLBO versus TLBOA	HLTLBO versus GBBO	HLTLBO versus PSO	HLTLBO versus CMA-ES
<i>fun</i> <sub>1</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>2</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>3</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>4</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>5</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>6</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>7</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>8</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>9</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>10</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>11</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>12</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>13</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>14</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>15</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>16</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>17</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>18</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>19</sub>	+ve	+ve	+ve	+ve
<i>fun</i> <sub>20</sub>	+ve	+ve	+ve	+ve

## 5 Conclusion

In this paper, half-life teaching factor is incorporated with teaching-learning-based optimization algorithm (TLBOA). This discovered algorithm is named as half-life teaching factor based TLBO (HLTLBO) algorithm. HLTLBO has been proposed for avoiding premature convergence of the algorithm. Half-life formula is incorporated into the teacher-phase (TP) of the TLBOA. By using 20 benchmark functions, the performance of HLTLBO has been evaluated. Comparison of the HLTLBO with TLBOA, global-best inspired biogeography-based optimization (GBBO), particle swarm optimization (PSO), and covariance matrix adaptation evolution strategy algorithm (CMA-ES) demonstrates that the HLTLBO is more competitive than other considered algorithms.



## References

1. Rao, R. V., Savsani, V. J., & Vakharia, D. P. (2011). Teaching–learning-based optimization: a novel method for constrained mechanical design optimization problems. *Computer-Aided Design*, 43(3), 303–315.
2. Singh, G., Sharma, N., Sharma, H. (2017). Intelligent neighbourhood teaching learning based optimization algorithm. In *2017 International conference on computing, communication and automation (ICCCA)* (pp. 986–991). Piscataway: IEEE.
3. Ghasemi, M., Ghanbarian, M.M., Ghavidel, S., Rahmani, S., Moghaddam, E.M. (2014). Modified teaching learning algorithm and double differential evolution algorithm for optimal reactive power dispatch problem: A comparative study. *Information Sciences*, 278, 231–249.
4. Kunjje, Y., Wang, X., & Wang, Z. (2016). An improved teaching-learning-based optimization algorithm for numerical and engineering optimization problems. *Journal of Intelligent Manufacturing*, 27(4), 831–843.
5. Zheng, S., & Ren, Z. (2016). Closed-loop teaching-learning-based optimization algorithm for global optimization. In *2016 12th world congress on intelligent control and automation (WCICA)* (pp. 2120–2125). Piscataway: IEEE.
6. Sharma, P. K., Sharma, H., & Sharma, N. (2016). Gbest inspired biogeography based optimization algorithm. In *IEEE international conference on power electronics, intelligent control and energy systems (ICPEICES)* (pp. 1–6). Piscataway: IEEE.
7. Kennedy, J. (2011). Particle swarm optimization. In *Encyclopedia of machine learning* (pp. 760–766). Berlin: Springer.
8. Chocat, R., Brevault, L., Balesdent, M., & Defoort, S. (2015). Modified covariance matrix adaptation-evolution strategy algorithm for constrained optimization under uncertainty, application to rocket design. *International Journal for Simulation and Multidisciplinary Design Optimization*, 6, A1.
9. Rao, R. V., Savsani, V. J., & Vakharia, D. P. (2012). Teaching–learning-based optimization: an optimization method for continuous non-linear large scale problems. *Information Sciences*, 183(1), 1–15.
10. Hansen, N. (2006). Compilation of results on the 2005 CEC benchmark function set. *Online*.
11. Tang, K., Yáo, X., Suganthan, P. N., MacNish, C., Chen, Y.-P., Chen, C.-M., & Yang, Z. (2007). Benchmark functions for the CEC'2008 special session and competition on large scale global optimization. *Nature inspired computation and applications laboratory, USTC, China* (Vol. 24).
12. Li, X., Tang, K., Omidvar, M. N., Yang, Z., Qin, K., & China, H. (2013). Benchmark functions for the CEC 2013 special session and competition on large-scale global optimization. *gene*, 7(33), 8.
13. Sharma, A., Sharma, H., Bhargava, A., Sharma, N., & Bansal, J. C. (2016). Optimal power flow analysis using lévy flight spider monkey optimisation algorithm. *International Journal of Artificial Intelligence and Soft Computing*, 5(4), 320–352.
14. Sharma, N., Sharma, H., Sharma, A., & Bansal, J. C. (2016). Modified artificial bee colony algorithm based on disruption operator. In *Proceedings of fifth international conference on soft computing for problem solving* (pp. 889–900). Berlin: Springer.

# Multilingual Data Analysis to Classify Sentiment Analysis for Tweets Using NLP and Classification Algorithm



Pragati Goel, Vikas Goel and Amit Kumar Gupta

**Abstract** The analysis of sentiments consists in identifying and classifying the opinions, attitudes, and sentiments of people expressed in original sentences. The advancement of social media, different critics, forum discussions, blogging, and working on social networks can be divided into different ways. Users who generate huge amounts of sentiment data on the website are in large quantities in the form of tweets and status updates. The sentiment analysis of this data is useful for market analysis and product research organizations. They are increasingly using public opinions in these media for their decision-making. In this paper, we propose an approach for analyzing the sentiment or opinion in an efficient manner. For this, we have proposed a technique that focuses multilingual data analysis to classify sentiment analysis for the tweets.

**Keywords** Opinion mining · Sentiment analysis · Applications · Opinionated data · Social media

## 1 Introduction

Today in this “Data Era”, there is a rapid increase in the amount of user generated data on social media platforms like Facebook, Twitter, LinkedIn, Instagram, and many more. Many opportunities and new open-door organizations have been motivated people for reviews about their products and services. That is why social media generates a lot of sentiments and abundant information from tweets, news,

---

P. Goel (✉)

Shri Venkateshwara University, Gajraula, UP, India  
e-mail: [goelpragati78@gmail.com](mailto:goelpragati78@gmail.com)

V. Goel

Ajay Kumar Garg Engineering College, Ghaziabad, UP, India  
e-mail: [rvikasgoel@gmail.com](mailto:rvikasgoel@gmail.com)

A. K. Gupta

KIET Group of Institutions, Ghaziabad, UP, India  
e-mail: [amitid29@gmail.com](mailto:amitid29@gmail.com)

blog postings, and more. Due to slang and misspellings, it is very difficult to compare Twitter's sentiment analysis with general emotional analysis. The maximum number of letters allowed by Twitter on a regional basis is 140. Various approaches such as knowledge-based approaches and machine learning approaches are available to analyze the emotions of sentences. In this article, we try to analyze Twitter posts (tweets) using an automatic learning approach. The analysis of emotions will probably identify the influence of information on emotional classification. We present a new feature vector to classify the polarity of tweets as Positive and Negative.

In this paper, we propose a method for opinion analysis. The study explains how to make the machines read and interpret the language that people use, i.e., natural language is NLP. Yet the word does not exist in the machine world. Machines represent words by the sequences of numbers, which are associated with a character while displaying them on screen. The Sentiment Analysis is the name of the problem in which the machine gets capability to analyze and predict with maximum precision possible the sentiment that will be obtained by a person. But as per our research, data analysis is one of the major critical problems for sentimental analysis. So we have proposed a technique that focuses multilingual data analysis to classify sentiment analysis for the tweets. Naïve Bayes classification algorithm is used for opinion analysis in the proposed technique.

The main objectives of this paper are as follows:

First, we want to figure out what classifiers and highlights deliver the best outcomes for this specific notion grouping assignment. We characterize the Twitter information by trying different things with a regulated classifier (i.e., Gullible Bayes).

Second, decide if posts are "certain", "negative", "both", "nonpartisan" or junk ("n/a"). Moreover, we likewise give an account of arrangement that comes about for posts labeled on the execution ("standard", "vernacular", or "n/an") and the investigations are accomplished in double phases: main phase utilizes highlight sets containing the most incessant expressions of the Stanford NLP, and another step utilizes include sets with the most useful expressions of the quantity, as estimated utilizing a data theoretic.

Third we speculate that multilingual information is harder to order (i.e., create bring down precision comes about). We aim to investigate and discover efficient algorithms and a method that may use to convert maximum data into once language so that we can fill the research gap. The most crucial one is multilingual sentimental analysis. This is mandatory because we have 6600 languages and every single opinion matters a lot. But tradition sentimental is only working on the English data.

Another theory is that the principal methodology (i.e., with highlight groups comprising of maximum continuous arguments) produces bring down outcomes than another methodology (i.e., through include sets encompassing the maximum enlightening words), since a continuous word is not really a significant component for assessment arrangement.

The rest of this paper is organized as follows. A related study explains the environment of research work in sentimental analysis. In the methodology section, the proposed scheme for sentimental analysis is presented. The performance of the

proposed technique is evaluated and analyzed in the performance measurement section. Finally, this article concludes by comparing the proposed technology in the conclusion section with existing technology, and the scope of the future will be described in the future research section. Related work discusses the surroundings of research work in sentiment analysis. In Methodology section, a proposed scheme for sentiment analysis is presented. The performance of the proposed technique is evaluated and analyzed in Performance Measure section. Finally, the paper is concluded with the comparison of proposed technique with existing technique in the Conclusion section and future scope is mentioned in the Future Work section.

## 2 Related Work

Sentiment Enquiry is the intensive exploration of how assessments and points of view can be identified with one's feeling and disposition appears in common dialect regard to an occasion. Late occasions demonstrate that the slant examination has come up to incredible accomplishment, which can outperform the positive versus negative and manage entire field of conduct and feelings for various groups and subjects. In the field of feeling examination utilizing distinctive methods great measure of research has been done for expectation of social assessments.

Christianini and Taylor published and shared knowledge about SVM, an automatic learning algorithm. The authors have succeeded in thoroughly explaining how to approach SVM algorithms and algorithms to implement them to solve practical problems [1].

Kopel et al. specified that it was essential to obtain an extreme value of extreme information by using an unclarified message. Similarly, the author argues that only progressive and destructive publications do not provide an adequate understanding of fair publications. Thinking about unbiased positions is a clear contrast between progressive messages and destructive messages. The authors found that one of the corpora containing most of the neutral documents did not give emotions that could be used as counters to test both the positive and negative results of the document [2].

Go et al. introduced a methodology for the automatic classification of emotions of Twitter messages. Each of the query term messages has been categorized as negative or positive. There, authors used remote monitoring to display feelings results on Twitter using the appliance's learning procedure. Maximum entropy, SVM, and Naive Bayes algorithm were applied to learning data, including emoticons, with an accuracy of over 80%. The authors also discussed pretreatment steps that have resulted in greater accuracy [3].

Pak et al. conduct phonetic examinations and design emotional classifiers to classify the positive, negative, and equitable findings of the archives. In order to train emotional classifiers, the authors proposed an approach to automatic collection of corpora. To analyze diffusion dissimilarities between neutral, negative and positive sets, we used Tree Tagger [4].

Tan et al. said that users shared similar opinions which are likely to be connected. The authors proposed the model that were generated from either by following the network that has been made by tagging different users with the help of “@” or by analyzing the network of Twitter follower/followed [5].

Geetika et al. proposed an understanding of the machine learning semantic inspection procedure to characterize sentence auditing and article auditing with respect to Twitter information. The important thing was to study many audits to use the currently named Twitter dataset. A simple by by strategy gives better results than maximum entropy and SVM. Semantic inspection when WordNet is followed by methodology, the accuracy has been increased from 88.2 to 89.9%. The training information index can be expanded to improve the process of recognizable proof of vector-related sentences, and WordNet can be extended to the study summary [6].

Abinash Tripathy et al. revealed that the opinion examination was the most undeniable branch of dialect preparation. The goal can be a type of (positive) or feedback (negative) feedback, as well as the connection of the outcomes gotten by applying Naive Bayes (NB) and transporter vector (SVM) position counts. These estimations are utilized to describe investigations prompting a positive or negative review. The datasets considered for the preparation and testing of their study designs are named according to the member film dataset and the correlation with the results available in the existing literature constitutes a critical examination [7].

Bac Le et al. investigated that twitter is a miniaturized scale blogging website in which clients can post refreshes (tweets) to companions (supporters). It has turned into an enormous dataset of the purported slants and acquaints an approach with determination of another list of capabilities. In view of Information Gain, Bigram, Object arranged extraction strategies in supposition investigation on long-range informal communication side additionally proposes an assumption examination display in view of Naive Bayes and Support Vector Machine. Its motivation is to break down conclusion all the more successfully [8].

Praveen Kumar et al. have depicted that a suitable conclusion examination can be performed on any zone by get-together, an example gathering of spectators evaluations from Twitter. Such inspections make valuable contributions to film associations and creators, TV action guidelines, and other things, and recognize that they add a negative influence to the way people feel about their area. By quickly discovering negative examples, they can learn about spectator satisfaction by relying on informed decisions about the most competent strategies targeting specific parts of their own. The calculations normally used for content game plans, such as Naive Bayes, SVM, decision trees, and random forests, have been redesigned. In evaluating these different counts, the authors found that Model J48 calculations had the highest analyzes of this dataset with 20 sporadic trees [9].

Bhavitha et al. center on the few machine learning procedures which are utilized as a part of breaking down the slants and in conclusion mining. The authors presented a detailed examination of the various machine learning procedures and subsequent contrasts, as well as their precision, points of interest and limitations of all methods. From research, they can deduce that managed learning strategies such as Naive Bayes and support vector machines are considered standard learning techniques. Vector

Machine Support offers incredible accuracy over various classifiers. The vocabulary-based methodology should be powerful because it requires a manual report. Larger entropy gets even better performance but is more experienced than fitting [10].

Shah et al. have proposed a new algorithm which can be called a hybrid algorithm. It uses three techniques for sentiment analysis. With the help of this movie reviews, product reviews, spam detection, and knowing consumer needs can be fulfilled [11].

Singh and Goel studied various machine learning algorithms for Twitter sentiment analysis and three-lexicon-based techniques for sentimental analysis. The study explores the different machine learning techniques in order to identify its usage and to increase interest in this research area [12, 13].

### 3 Proposed Approach

Twitter is a multilingual data source. This feature of Twitter is not used in the previous works and only English language tweets have been considered in the dataset. We have developed a language-independent system which will enhance the dataset by converting multilingual tweets into English language with the help of Google Translator API.

Second the tweets were in the unstructured textual format and were converted to meaningful data by using the preprocessing step. This processed data is further worked upon and converted to numerical vectors using dictionary modeling and feature extraction. On this labeled dataset set finally, Naïve Bayes classifier is applied and confusion matrix is generated to measure the performance.

### 4 Methodology

During the literature survey, we observe that many authors just want to increase the accuracy by using different algorithms (Naive Bayes, max entropy, and SVM and Decision tree). We observe that Twitter is the source of many language communications but existing system can only have facility of English data analysis.

- (a) This is the first improvement of the system. We are trying to develop language-independent system. This is done using Google translator API.
- (b) We apply all preprocessing step in order to filter dataset.
- (c) We apply Stanford NLP for data modeling and training.
- (d) We apply Naïve Bayes for machine learning.

Every author tried to increase the accuracy by following the tradition way. In the tradition way, we train the machine by limited number of sentence. But there is serious problem occur that machine knowledge is limited because of having limited data training. So in our approach we proposed a solution with Stanford NLP which

is the modeling of English language. This library is having very good knowledge and all possible combination for training.

This system proposes to develop functions for the users so that they do not encounter problems when they use missing data, one-way contacts, one-way contacts, etc. Because we collect data on Twitter to develop this project, users will not be bothered by search, which is based on keywords. As we maintain a feature extraction method that generates a generic output, we can directly implement a classification method. The feasibility study is a complete process of analysis and complete system design. The search begins with the classification of the problem definitions. The feasibility is to judge whether it is worth it or not. The analyst develops a logical model of the system only when an accepted definition of the problem is generated. Research alternatives are carefully analyzed. Figure 1 shows the flow of our proposed work and the various steps of our proposed work.

*Naive Bayes (NB) Classifier:* This classifier is based upon the feature theorem of Bayes and is a probabilistic classifier. It makes use of the properties of Bayes theorem assuming the strong independence between the features or characteristics. One of the upside of this classifier is that it requires little measure of preparing information to ascertain the parameters for forecast. Rather than calculating the complete covariance matrix, best variance of the characteristics is computed due to independence of features.

The conditional probability is  $P(e|a)$  for each class “a”(positive, negative) given a literary evaluation or review “e”. The following equation can be used to compute this value according to Bayes theorem:

$$P(a|e) = P(e|a) * P(a)/P(e) \quad (1)$$

To further compute the term  $P(e|a)$ , the break up is expressed in the following equation where it is assumed that  $f_i$ 's are conditionally independent given  $e$ 's class.

$$P_{NB}(a|e) = P(a) \left( \sum P(f_i|a) \right)^{ni(e)} / (P(e)) \quad (2)$$

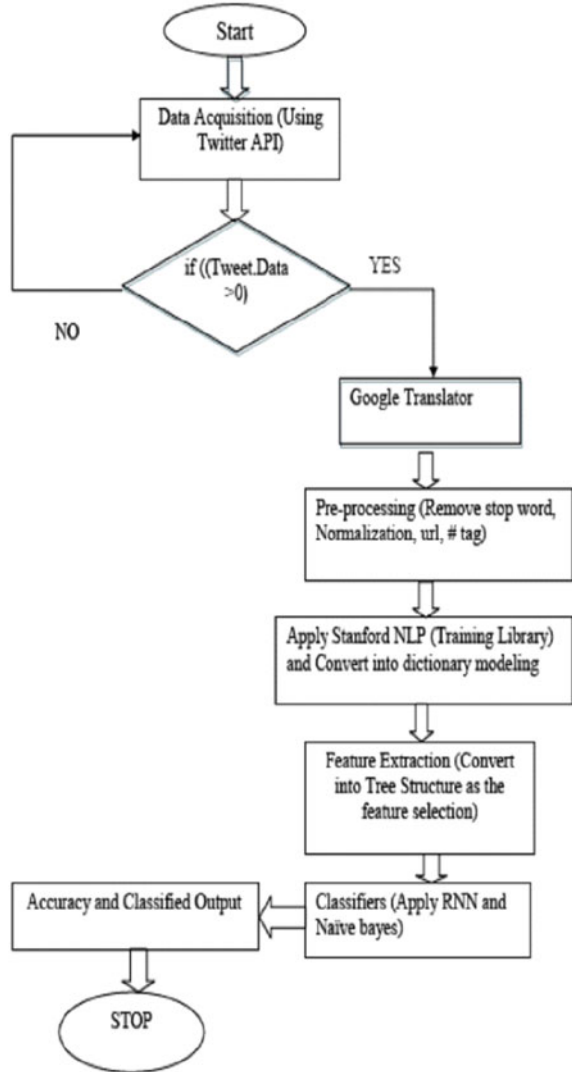
In this formula,  $f$  represents a feature and  $ni(e)$  represents the count of feature  $f_i$  found in tweet  $e$ . There are a total of  $m$  features. Parameters  $P(a)$  and  $P(f_i|a)$  are obtained through maximum likelihood estimates.

## 5 Performance Measure

The opinion classification can be done using the following equations for precision and accuracy which are evaluated using four indexes [7].

*Precision:* The classifier's precision value is evaluated here which is the proportion of number of accurately anticipated positive surveys to the aggregate number of audits anticipated as positive.

**Fig. 1** Flow diagram of proposed work



$$\text{Precision} = \text{tp} / (\text{tp} + \text{fp}) \tag{3}$$

*Accuracy*: It is one of the utmost communal presentation calculation parameter and it is intended as the percentage of number of properly expected analyzes to the aggregate number of reviews present in the quantity. The formula for calculating accuracy is given as follows:

$$\text{Accuracy} = (\text{tp} + \text{tn}) / (\text{tp} + \text{tn} + \text{fp} + \text{fn}) \tag{4}$$



**Table 1** Result of parameters for Naive Bayes

Algorithm used	Dataset	Precision	Accuracy %
Naïve Bayes	2000	0.78	79.4

**Table 2** Precision and accuracy for Naive Bayes Classifier

Author	Twitter dataset	Precision	Accuracy
Proposed	2000	0.78	79.4
Geetika Gautam et al.	2000	0.44	88.6
Hailong Zhang et al.	2000	0.65	68.75

- True Positive (tp) which is correctly classified as positive.
- False Positive (fp) which is incorrectly classified as positive
- True Negative (tn) which is correctly classified as negative.
- False Negative (fn) which is incorrectly classified as negative.

Table 1 shows parametric result of Naïve Bayes classifier with respect to accuracy and precision with dataset of 2000. Accuracy has been achieved at 79.4 and Precision 0.78.

## 6 Comparative Analysis

In this section, the comparative analysis of result obtained by our proposed method for the same dataset size with the output obtained in other research papers is done. Also Naïve Bayes classifier is applied on the dataset in all the manuscripts.

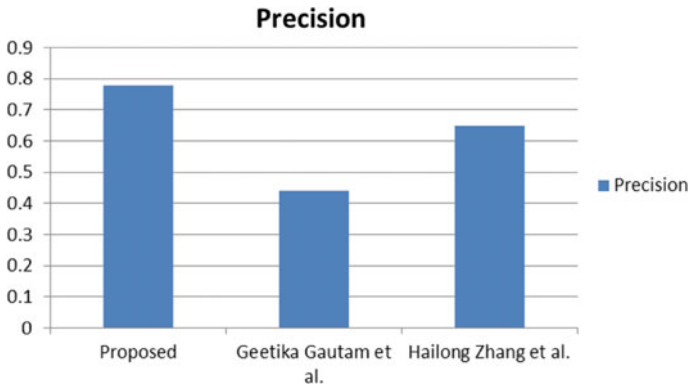
Various authors have worked on Naive Bayes (NB) classifier. We compare our work with Geetika et al. [6] and Zhang et al. [14] who had also taken the same dataset of values.

Table 2 depicts work of Geetika Gautam et al. has high accuracy, as compared to others. Table 2 also depicts that our proposed method has better precision value as compared to the other two research papers.

Figure 2 shows the bar graph for precision as per the values of Table 2 depicting highest precision for the proposed work.

## 7 Conclusion

In our examination, we have made an effort to perform sentiment analysis for conclusion investigation for the tweets as negative or positive. We have brought information from Twitter utilizing machine learning strategies.



**Fig. 2** Graph of precision for Naive Bayes Classifier

We have implemented Naïve Bayes (NB) algorithm and the results show that our proposed algorithm has outperformed in precision factor. In case of accuracy, our proposed algorithm is better than Hailong Zhang et al. work but less than Geetika Gautam et al. work. This may be due to single language and multilingual tweets. This may be further improved in future. Through the support of outcomes we can presume that machine learning calculation is finest for arrangement of assumption examination.

## 8 Future Work

Twitter has a limit of 140 characters per tweet and is used by a large number of people to express their views so it provides result of better quality. Other than Twitter, we will intend to expand our research work for other social media platforms too like Facebook, Instagram, etc. Also, other steps can be included in preprocessing like considering emoticons and domain name of URL, etc. If quality of data is improved, classification algorithms will be able to produce better quality of results. Future work will involve investigation of other approaches for preprocessing tweets because they have to be more thoroughly filtered to achieve the higher accuracy, precision, etc.

Another experiment that may be carried out is the replacement of the abbreviations with their full meaning. It obviously will increase the size of the training corpus but may add more sense to the tweet.

## References

1. Christianini, N., & Taylor, J. S. (2001). An introduction to support vector machines and other kernel-based learning methods (Vol. 30(1), pp. 103–115). Cambridge University Press, Kybernetes.
2. Koppel, M., & Schler, J. (2006). The importance of neutral examples for learning sentiment. *Computational Intelligence*, 22(2), 100–109.
3. Bhayani, R., & Huang, L. (2009). *Twitter sentiment classification using distant supervision*. CS224 N Project Report, Stanford (Vol. 1, pp. 12–21).
4. Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. *LREc*, 10, 1320–1326.
5. Tan, L. L., Tang, J., Jiang, L. et al. (2011). User-level sentiment analysis incorporating social networks. In *17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1397–1405).
6. Gautam, G., & Yadav, D. (2014). Sentiment analysis of twitter data using machine learning approaches and semantic analysis. In *Conference on Contemporary Computing (IC3) IEEE* (pp. 437–442).
7. Tripathy, A., Agrawal, A., & Rath, S. K. (2015). Classification of sentimental reviews using machine learning techniques. *International Conference on Recent Trends in Computing (ICRTC)*, 57, 821–829.
8. Le, B., & Nguyen, H. (2015). Twitter sentiment analysis using machine learning techniques. In *Advanced Computational Methods for Knowledge Engineering* (pp. 279–289). Springer International Publishing Switzerland.
9. Kumar, P., Choudhury, T., Rawat, S., & Jayaraman, S. (2016). Analysis of various machine learning algorithms for enhanced opinion mining using twitter data streams. In: *International Conference on Micro-Electronics and Telecommunication Engineering IEEE* (pp. 265–270).
10. Bhavitha, B. K., Rodrigues, A. P., & Chiplunkar, N. N. (2017). Comparative study of machine learning techniques in sentimental analysis. In: *International Conference on Inventive Communication and Computational Technologies (ICICCT) IEEE* (pp. 216–221).
11. Shah, S., Kumar, K., & Sarvananguru Ra, K. (2016). Sentimental analysis of twitter data using classifier algorithms. *International Journal of Electrical and Computer Engineering*, 6(1), 357.
12. Singh, R., & Goel, V. (2017). Various machine learning algorithms for twitter sentiment analysis. In: *Third International Conference on Information and Communication Technology for Competitive Strategies (ICTCS)* (pp. 1–10). Springer proceedings (LNNS).
13. Singh, R., & Goel, V. (2017). Comparative study for sentiment analysis of twitter data. *International Journal for Research in Applied Science & Engineering Technology*, 5(XII), 1485–1491.
14. Zhang, H., Gan, W., & Jiang, B. (2014). Machine learning and lexicon based methods for sentiment classification: A survey. In: *Conference on Web Information System and Application (WISA) IEEE* (pp. 262–265).

# **Intelligent Hardware and Software Design**

# Pedestrian–Autonomous Vehicles Interaction Challenges: A Survey and a Solution to Pedestrian Intent Identification



Pranav Pandey and Jagannath V. Aghav

**Abstract** Autonomous Vehicles are on rise around the globe, millions of them are already there on road with medium levels of automation but still there is a long way to go for full autonomy. One of the biggest roadblocks for autonomous vehicles to reach full autonomy is driving in urban environments. To make autonomous vehicles fully autonomous, they require the ability to communicate with other road users (pedestrian, vehicles, and other road users) and understand their intentions. Social interaction is a complex task, there are uncountable scenarios that happen on roads that require human interaction both verbal and nonverbal. Deciding whether a person standing on the sidewalk is about to cross the road, or they are just waiting near the sidewalk is a difficult task for an autonomous vehicle, and it could be a matter of life-and-death in case of a vehicle driving at very high speed. So, it is very important for self-driving cars to identify true intentions of on-road pedestrians and understand social interaction norms. In this paper, we go through some of the challenges in Pedestrian and Autonomous vehicles interaction that autonomous vehicles might face while driving in an urban environment; after that we propose a novel architecture for identifying pedestrian's intention using pedestrian's detection, pose estimation, and classification algorithms while discussing different methods of each.

**Keywords** Self-driving cars · Machine learning · LIDAR · Automotive

## 1 Introduction

Fully automated vehicles are still a long way to go and one of the biggest roadblocks we see today is to teach them social interaction. Social interaction plays an important role in resolving various potential ambiguities in traffic and to drive properly in real world. There are a lot of challenges that needs to be addressed to create a system

---

P. Pandey (✉) · J. V. Aghav  
College of Engineering, Wellesley Rd, Shivajinagar, Pune 411005, Maharashtra, India  
e-mail: [pranavp17.comp@coep.ac.in](mailto:pranavp17.comp@coep.ac.in)

J. V. Aghav  
e-mail: [jva.comp@coep.ac.in](mailto:jva.comp@coep.ac.in)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_27](https://doi.org/10.1007/978-981-15-0694-9_27)

that can interact socially [1]. Road crashes kill about 1.3 million people around the globe every year and the count of severely injured people is at around an estimated 50 million [2]; this count has been increasing continuously for last few years. The main road users that are under concern are—pedestrians, cyclists, and motorcyclists because they travel on foot or on light vehicles which make them vulnerable to accidents and heavy damage. Avoiding vehicle to pedestrian crashes on road is one of the most important aspect of self-driving cars and toward that end there is already a huge on-going research in field of detecting pedestrians on road—but that is only one part of the solution, to efficiently know where the pedestrians are on the road and where they will be going in future as the vehicle travels is the key to avoid fatalities and for that we also need to identify the intention of the pedestrians on road and know if someone is about to come in the path of the self-driving car or not. Given that autonomous vehicles may commute without any passengers on board, they are subject to malicious behavior, similar to those observed against a number of autonomous robots used in malls. Understanding the true intention of these people can help the vehicles act accordingly [3], and then take the precautionary measures. In this paper, we will first briefly look into how an autonomous vehicle works, and then we will see some challenges in pedestrian–autonomous vehicles interaction, and toward the end we will propose a solution to identify pedestrian intent in real time.

## 2 Challenges

### 2.1 *Gesture Recognition*

Humans are very quick in recognizing other people's gestures and behaviors on road and can take action accordingly, and autonomous vehicles need about same level of this ability to run easily on roads [4]. There are subtle signals that humans take for granted: the body language of a traffic control officer, for example, or a bicyclist trying to make eye contact. The challenge here is that how do you teach a car to understand spoken commands and hand signals from law enforcement employees that it has not seen ever before or dealt with gesture from other drivers, how do you teach a computer–human intuition? Perhaps, endless road training is the only way to teach the vehicles about social interaction among humans. This is one of the biggest challenge in autonomous vehicles today, and we will discuss it in detail about this issue later in this paper.

## ***2.2 Reliably Recognizing Traffic Norms Where the Traffic Signals Are not Working***

With the advancement of computer vision, Autonomous vehicles can now recognize traffic lights reliably. But the problem starts in the case of a power failure or places where traffic lights are malfunctioned, the vehicle should be able to correctly interact with other vehicles drivers and follow the traffic norms to run smoothly. Here again, it is a question of human autonomous vehicles interaction—how to teach autonomous vehicles to interact with other human driving the vehicles and run properly on these types of intersections.

## ***2.3 Lane Cutting and Making Turns to Join Roads with Fast-Moving Traffic***

Think about cars merging from a side road onto a busy highway with toll both. Humans are good at this task of entering into a new road lane, they know that by making eye contact only they can assert their need and the other driver will give them way to merge from the side road onto the highway, and all of this decisive interaction takes place in a split of a second. How exactly should people have this interaction with a self-driving car? It is something that has yet to be established.

## ***2.4 Judgment Calls***

Making judgment calls is a difficult task even for humans and a wrong call can affect many lives. Sometimes drivers face scenarios where when they have to come to a sudden halt due to some obstacle on road which can be programmed easily, but the difficult task is to program the car what to do if there is some obstacle on road and it is a busy highway—should it stop and let other cars collide from back and cause a pileup of cars or should it hit the obstacle; the autonomous vehicles will take a lot of time taking this type of decision based on the likelihood of both the scenarios. What is more, self-driving vehicles will not necessarily be able to decide between swerving and hitting a child on the road and hitting a pile of junk that is blocking the lane.

## ***2.5 Kangaroos on Road***

Apart from humans, the autonomous vehicles also need to interact well with the animals all around as they are very common on roads around the globe, and among

animals there is a very curious case in Australia where KANGAROOS create a lot of problems for autonomous vehicle testing as conducted by Volvo. According to Volvo the kangaroos jump from one place to another for movement and while they are in the air it is difficult for vision system of the autonomous vehicles to detect the correct distance of the kangaroo as they appear to be far away because they are in air, and when they land back they suddenly appear to be very close to the car which confuses the system and can cause problems in life-and-death situations.

### 3 Dealing with Pedestrians

#### 3.1 Pedestrian Detection

Pedestrian detection is a challenging task because pedestrians appear in different types of clothing which might not be easily identifiable, they might have different possible articulations, also there can be some occlusion due to accessories the pedestrian uses or other things present on roads [5], sometimes in crowded paces there is also frequent occlusion on pedestrians between themselves which makes it a difficult task to identify pedestrians correctly [6] in the real world by self-driving cars [7, 8].

In the figure (Fig. 1) we see a typical flowchart of how an autonomous vehicle identify whether an object present in the image obtained by the camera is a pedestrian or not [9, 10]. Pedestrian detection is one of the well-researched domains in the field

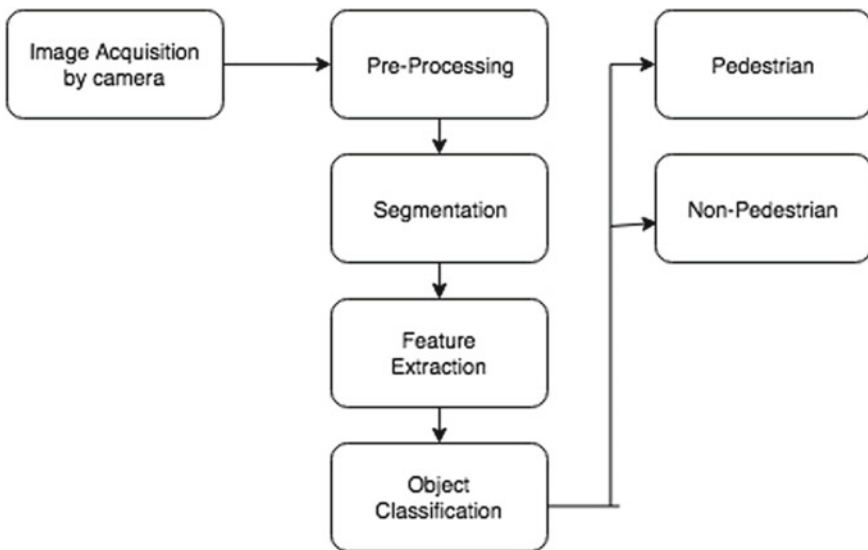
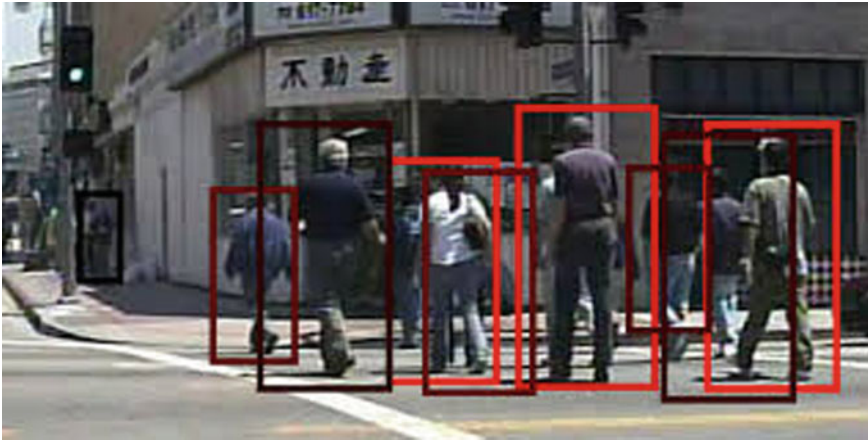


Fig. 1 Flowchart of typical pedestrians





**Fig. 2** Pedestrian detected in INRIA dataset [25]

of autonomous vehicles [11] and there are many different approaches followed by different big players in the fields of self-driving car—like Waymo (formerly Google self-driving car project), Uber, Daimler, and General Motors (Fig. 2).

From the above-stated challenges, it is clear that it is very important to provide autonomous vehicles with human intuition so that the vehicles can make decision and understand other humans on foot or any other vehicle.

After studying Pedestrian Detection literature, we believe that it can be classified into three major families [12, 13, 14]:

- (1) Deformable Parts Model(DPM) [15]
- (2) Deep Networks
- (3) Decision Forests

From Fig. 3 we can see that Decision Forest Model reach best performance in the task of pedestrian detection [16] (Note: Most of the models are trained on INRIA and Caltech-USA dataset, but we plan to train our model on Daimler dataset).

### **3.2 Pedestrian Intent Identification**

Pedestrian intent identification is one of the most important and the most challenging task for a self-driving vehicle to address, as it can make a huge difference for the vehicles in becoming a reality and coming on roads at a large scale. In the literature we have studied [3] most of the work in pedestrian intent identification is based on two works—using Kalman Filters [17] and Head Pose Estimation with Pedestrian movement dynamics [18, 19]; Kalman filtering, also known as Linear Quadratic Estimation (LQE), is an algorithm that uses a series of measurements observed over time, containing statistical noise and other inaccuracies, and produces estimates

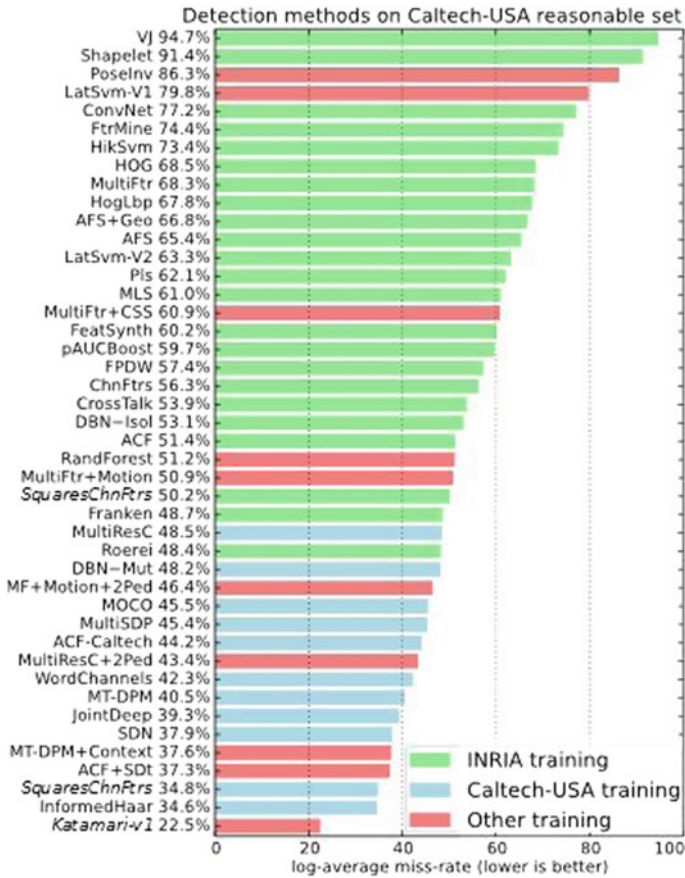


Fig. 3 Different pedestrian detection algorithms performance

of unknown variables that tend to be more accurate than those based on a single measurement alone, by estimating a joint probability distribution over the variables for each timeframe [20].

Other commonly used method is using Head Pose Estimation and Pedestrian Dynamics like position and velocity for path prediction [21]. Also, most existing solutions work on stereo vision cameras, which require high processing power and response time is high. In our system, we will use Daimler Pedestrian Path prediction dataset which is also a stereo vision dataset, but we will be using photos from only one side of the camera [22, 23] and design a monocular vision-based system which will be robust and require less processing power.

## 4 Proposed System

### 4.1 Aim

The aim of this proposed system is to create a novel architecture for pedestrian intent identification and path prediction which can be applied to autonomous vehicles and plethora of other instances, it will help in avoiding accidents by correctly predicting the pedestrians path and checking the chances of the pedestrian getting into the path of the autonomous vehicle and taking precautionary measures to avoid any kind of catastrophe that could arise due to careless behavior of the pedestrian or driver, driving under the influence of alcohol and other drugs or any other similar situation.

### 4.2 System Design

We use pedestrian detection algorithms to first identify the pedestrians present in the image taken by the vehicle and make a bounding box around the pedestrians, then we will apply pose estimation algorithms to identify the pose of the pedestrian by placing a skeleton figure on the pedestrian present in the image. After getting the skeleton figure, we extract features from that figure that are essential for the next step. After this we run classification algorithms, which will classify the pedestrian in one of the four classes present, i.e., crossing, stopping, bending, and starting; this will be our final output. As the output, we get the correct pedestrian intention, classified as one among four—stopping, crossing, bending, and starting which will help the vehicle to take decision instantly and avoid any bad situation.

### 4.3 Architecture

The proposed architecture consists of three main modules

- (1) Pedestrian Detection
- (2) Pose Estimation
- (3) Pedestrian Intent Classification

First the camera takes video of the real world and the video is divided into frames and then each frame is sent to pedestrians detection module where the image is processed and then the pedestrian detection algorithm identifies the pedestrian present in the image and bounding boxes are created around the pedestrians, then the image with bounding boxes around the pedestrians is feed to the next module where the pedestrians present in the image are analyzed by pose estimation algorithm to identify them and place skeleton system on the pedestrians for their exact pose estimation, then finally this skeleton structure containing image is fed to the last module where

we use classification algorithm to classify the detected pose among the four classes that we have used for training to correctly identify the pedestrians intentions.



Steps for identifying pedestrians' intentions

#### 4.4 Algorithm

Result: Pedestrian Intention: crossing, stopping, bending, or starting.

Initialization;

**While** *driving* do:

Take video;

Divide in Frames;

Detect Pedestrians;

Estimate Pose;

Classify Pose is one of the four classes;

**End**

## 5 Conclusion

In the literature, different algorithms used for Pedestrian Detection and Intent Identification are discussed. We have seen that pedestrian detection is a well-researched area and there are algorithms which perform very well using decision forests based model and also DPM and deep network models, so will be using those existing models only for pedestrian detection, then we will use pose estimation algorithms—DensePose which is state-of-the-art pose estimation algorithm and try to fit skeleton models on detected pedestrians. We will try to design algorithm which will work on monocular vision as most existing solutions work on stereo vision cameras only as discussed above which requires high processing power and due to which response time is high, which gives us a very less distance for safety response. In our proposed system, we will use Daimler Pedestrian Path prediction dataset [24] which is also a stereo vision dataset, but we will be using photos from only one side of the camera and design a monocular vision-based system which will be robust and require less processing power and then try to use different types of classification algorithms to attain high accuracy. In future research works we will try to classify the pedestrian intents in many other classes rather just only in the four classes we do here, as pedestrian show complex behaviors on road and cannot be always among the four classes we have used here.

## References

1. Mahadevan, K., Somanath, S., & Sharlin, E. (2018). Communicating awareness and intent in autonomous vehicle–pedestrian interaction. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems—CH'18*. <https://doi.org/10.1145/3173574.3174003>.
2. Yan, J., Lei, Z., Wen, L., & Li, S.Z. (2014). The fastest deformable part model for object detection. In *CVPR*.
3. Völz, B., Behrendt, K., Mielenz, H., Gilitschenski, I., Siegwart, R., & Nieto, J. (2016). A data-driven approach for pedestrian intention estimation. In *Proceedings of the IEEE International Conference on Intelligent Transportation Systems (ITSC), Rio de Janeiro, Brazil*, 1–4 November 2016.
4. Rasouli, A., Kotseruba, I., Tsotsos, J. (2017). Agreeing to cross: How drivers and pedestrians communicate. [arXiv:1702.03555v1](https://arxiv.org/abs/1702.03555v1).
5. Xu, X., & Fan, C. (2018). Autonomous vehicles, risk perceptions and insurance demand: An individual survey in China. *Transportation Research Part A: Policy and Practice*. <https://doi.org/10.1016/j.tra.2018.04.009>.
6. Keller, C., & Gavrila, D. (2014). Will the pedestrian cross? A study on pedestrian path prediction. *IEEE Transactions on Intelligent Transportation Systems*, 15(494–506), 9.
7. WHO Survey on Road Safety-Mortality rate. Retrieved from <https://www.who.int/gho/roadsafety/mortality/en/>.
8. Pinheiro, P., & Collobert, R. (2014). Recurrent convolutional neural networks for scene labeling. In: *JMLR*.
9. Brunetti, A., Buongiorno, D., Trotta, G. F., & Bevilacqua, V. (2018). Computer vision and deep learning techniques for pedestrian detection and tracking: A survey. *Neurocomputing*, 300, 17–33. <https://doi.org/10.1016/j.neucom.2018.01.092>.
10. Mogelmoose, A., Trivedi, M. M., & Moeslund, T. B. (2015). Trajectory analysis and prediction for improved pedestrian safety: Integrated framework and evaluations. In *2015 IEEE Intelligent Vehicles Symposium (IV)*. <https://doi.org/10.1109/ivs.2015.7225707>.
11. Cao, Z., et al. (2017). Realtime multi-person 2D pose estimation using part affinity fields. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. <https://doi.org/10.1109/cvpr.2017.143>.
12. Benenson, R., Omran, M., Hosang, J., & Schiele, B. (2014). Ten years of pedestrian detection, what have we learned? *ECCV 2014: Computer Vision—ECCV 2014 Workshops* (pp. 613–627).
13. Enzweiler, M., & Gavrila, D. (2009). Monocular pedestrian detection: Survey and experiments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31, 2179–2195.
14. Dollar, P., Wojek, C., Schiele, B., & Perona, P. (2012). Pedestrian detection: An evaluation of the state of the art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(4), 743–761. <https://doi.org/10.1109/tpami.2011.155>.
15. Cao, Z., Simon, T., Wei, S., & Sheikh, Y. (2017). Realtime multiperson 2D pose estimation using part affinity fields. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. <https://doi.org/10.1109/cvpr.2017.143>.
16. Zhang, S., Benenson, R., Omran, M., Hosang, J., & Schiele, B. (2017). Towards reaching human performance in pedestrian detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. <https://doi.org/10.1109/tpami.2017.2700460>.
17. Patel, H. A. Thakore, D. G. (2013). Moving object tracking using Kalman Filter. *International Journal of Computer Science and Mobile Computing*.
18. Rosenfeld, A., Zemel, R., & Tsotsos, J. K. The elephant in the room.
19. Cho, H., Rybski, P. E., Bar-Hillel, A., & Zhang, W. (2012). Real-time pedestrian detection with deformable part models. In *2012 IEEE Intelligent Vehicles Symposium*. <https://doi.org/10.1109/ivs.2012.6232264>.
20. Wojek, C., Walk, S., & Schiele, B. (2009). Multi-cue on-board pedestrian detection. In: *CVPR*.

21. Rasouli, A., Kotseruba, I., & Tsotsos, J. K. (2017). Are they going to cross? A benchmark dataset and baseline for pedestrian cross-walk behavior. In *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*. <https://doi.org/10.1109/iccvw.2017.33>.
22. Fang, Z., Va'zquez, D., & Lo'pez, A. (2017). On-Board detection of pedestrian intentions. *Sensors*, *17*(10), 2193. <https://doi.org/10.3390/s17102193>.
23. Ratsamee, P., Mae, Y., Ohara, K., Takubo, T., & Arai, T. (2012). People tracking with body pose estimation for human path prediction. In *2012 IEEE International Conference on Mechatronics and Automation*. <https://doi.org/10.1109/icma.2012.6285114>.
24. Schneider, N., Gavrilu, D. M. (2013). Pedestrian path prediction with recursive Bayesian Filters: A comparative study. *Lecture notes in computer science pattern recognition* (pp. 174–183). <https://doi.org/10.1007/978-3-642-40602-718>.
25. Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*.
26. Schulz, A. T., & Stiefelhagen, R. (2015). Pedestrian intention recognition using Latent-dynamic Conditional Random Fields. In *2015 IEEE Intelligent Vehicles Symposium (IV)*. <https://doi.org/10.1109/ivs.2015.7225754>.

# Code Profiling Analysis of Rough Set Theory on DSP and Embedded Processors for IoT Application



Vanita Agarwal, Rajendrakumar A. Patil and Jyoti Adwani

**Abstract** Rough set theory is a powerful artificial intelligence based tool used for data analysis and mining inconsistent information systems. In the presence of inconsistent, incomplete, imprecise, or vague data, normal statistical based data analytic techniques lag behind. This paper discusses the code profiling for rough set theory on DSP and ARM processors. This work was undertaken to understand the performance of rough set theory on existing processors for mining/analyzing inconsistent nature of IoT application at fog/edge interface.

**Keywords** Rough set theory (RST) · Internet of things(IoT) · Inconsistent information systems (IIS)

## 1 Introduction

The Internet of Things (IoT) architecture comprises of devices at cloud, fog, and edge interfaces for computing and data analytic. The cloud platform can work with hardware platforms with huge computational requirements, processing capabilities, workloads, etc. Unlike cloud based devices, the fog and specially the edge interface along with the sensor networks is limited in computational capabilities, processing capabilities with an emphasis on power/energy consumption when dealing with inconsistent information systems in IoT applications.

The authors have undertaken the code profiling analysis, with an objective to understand the performance of data analytic for inconsistent nature of IoT data on

---

V. Agarwal (✉) · R. A. Patil  
Electronics and Telecommunication Department, College of Engineering Pune, Pune,  
Maharashtra, India  
e-mail: [vsa.extc@coep.ac.in](mailto:vsa.extc@coep.ac.in)

R. A. Patil  
e-mail: [rap.extc@coep.ac.in](mailto:rap.extc@coep.ac.in)

J. Adwani  
Centre for VLSI and Nanotechnology, VNIT, Nagpur, Maharashtra, India

various hardware platforms used in IoT infrastructure. The IoT based smart system is diverse with Arduino/PIC/ARM /Intel Galileo as potential candidates for 8/16/32 bit processors.

The reminder of this paper is organized as follows: Sect. 2 highlights the existing literature about Rough Set Theory (RST). This section also shows that rough set theory has been used by several researchers in the past for analyzing inconsistent data. Section 3 discusses the experimentation done for the code profiling of RST on DSP and ARM processors. This section also highlights the frequency counts of occurring of various assembly level instructions for RST on DSP and ARM processors. Section 4 highlights the generated rule sets from different applications of IoT. At the end, the paper concludes by justifying a need for developing specific instruction set for rough set theory.

## 2 Related Work

An inconsistent information system signifies inconsistent, incomplete, vague, and/or imprecise data. In 1982, Prof. Pawlak Z. proposed Rough Set Theory (RST) [1–3] for reasoning/mining inconsistent data by evaluating equivalence relations between two sets of inconsistent data and partitioning it on the basis of concepts generated. Since then, RST has been used in a variety of applications for machine learning, decision analysis, data analytic, data mining, patter recognition [4–10]. For computation purposes of Rough Set Theory (RST), various software such as ROSE [11], RSES, WEKA, ROSETTA, Rough Sets [5, 12] are available which can be used on x86 processors.

Rough set theory can also be considered as a potential candidate for analyzing inconsistency in IoT applications at fog/edge interfaces due to its specialized features for missing data, redundancy isolation for data storage, and information overloading for the next generation IoT hardware exploration [13].

## 3 Experimentation

As IoT environment is usually non processor specific and many types of hand held portable gadgets are used, C language was used for defining the source code for RST. Various constructs of RST such as Elementary Set (ES), Crisp Set (CS), Lower Approximation  $L_A$ , Upper Approximation  $U_A$ , Core, and Reduct have been designed using C language. Our C code is different from others and can be completely implemented on hardware. **Data structures such as linked list and others were completely avoided.** The inconsistent data set considered for our experimentation is shown below in Table 1. The % Coal, sulfur, and phosphorus are the conditional attributes and Crack found is a decision attribute. The objects of study are the six



**Table 1** Inconsistent information table for quality monitoring of pipes in a factory environment

Pipes	% Coal	% Sulfur	% Phosphorus	Crack found
1	High	High	Low	Yes
2	Avg	High	Low	No
3	Avg	High	Low	Yes
4	Avg	Low	Low	No
5	Avg	Low	High	No
6	High	Low	High	Yes

pipes as shown. **An information system is called inconsistent when the same condition attribute values lead to different concepts.**

The code was compiled and debugged and output was generated using GCC compiler. The outputs match with the ROSE software outputs. The codes were also run on DSP and embedded hardware platforms like DSP TMS320 C6713 and ARM Cortex M4 boards, respectively. Code Profiling was also performed on both the boards [14].

### 3.1 Code Profiling of RST Algorithm on DSP and ARM Processor

The code profiling exercise was undertaken on TMS320C6713 (TI DSP processor) development board using CCS 5.1 and STM32F411 (ARM Cortex M4F) development board using CooCoX IDE [14]. Various inconsistent data sets were considered of varying number of conditional and decision attributes and varying number of samples. Table 2 shows the code profiling results of one such sample data set for integer and float data. The data set under consideration consisted of only six objects (i.e., Universe  $U$  of elements). Version 1 basically had integer data and version 2 had float data.

The frequency of occurring of assembly level instructions for elementary set, crisp set,  $L_A$ , and  $U_A$  (Rough Set Technique Constructs) on DSP TMS320C67X have been mentioned in Tables 2 and 3. Figures 1 and 2 show the plot of frequency vs assembly instructions generated for various rough set technique constructs (Reduct, Boundary, Accuracy, and Core, respectively) on DSP TMS320C67X.

The frequency of occurring of assembly level instructions for elementary set, crisp set,  $L_A$ , and  $U_A$  (Rough Set Technique Constructs) on ARM Cortex M4 have been mentioned in Tables 4 and 5. Figures 3 and 4 show the plot of frequency vs assembly instructions generated for various rough set technique constructs (Reduct, Boundary, Accuracy, and Core, respectively) on ARM Cortex M4.

The assembly instructions generated for various rough set constructs were tabulated separately in four categories .L, .M, .S, and .D as shown in Table 6 to further classify them as data processing/load store type/branch instructions.

**Table 2** Frequency count of assembly instructions generated for elementary set on DSP TMS320C67x

Instruction	Version 1	Version 2
ADD.L1X	9	4
MV.L1X	17	13
ZERO.L1	1	1
ADD.L1	–	4
ZERO.L2	5	3
CMPGT.L2	10	8
ADD.L2	18	10
MV.L2	17	13
CMPEQ.L2	7	15
MV.L2X	2	2
ADD.L2X	–	2
MVK.S1	8	8
MVKH.S1	8	8
SHL.S1X	8	6
MV.S1X	–	2
B.S1	–	25
B.S2	–	3
MVK.S2	30	24
MVKH.S2	26	22
LDW.D1	35	20
STW.D1	33	10
ADDAD.D1	6	6
ADDAW.D1	6	–
STW.D2	105	69
LDW.D2	74	51
ADDAW.D2	17	–
ADDAD.D2	6	4
NOP	181	186

This study has helped the authors in identifying the instructions for analyzing inconsistent data. This study also gave a preliminary understanding of the need for designing new and dedicated instruction sets for the support of rough set theory.

**Table 3** Frequency count of assembly instructions generated for  $L_A, U_A$  on DSP TMS320C67x

Instruction	Version 1	Version 2
ZERO.L1	3	4
SHL.L1X	–	1
MV.L1X	3	7
NOP	250	389
ADD.L1	1	1
CMPEQ.L1X	1	–
ZERO.L2	11	13
CMPGT.L2	8	8
INTSP.L2	–	2
ADD.L2	21	19
MV.L2	15	9
CMPEQ.L2	13	11
XOR.L2	6	6
MV.L2X	3	3
ADD.L2X	2	2
CMPLT.L2	13	10
B.S1	49	49
MVK.S1	4	4
MVKH.S1	4	5
CMPEQ.S1X		2
ADDK.S1	–	3
SHL.S1X	1	–
MVK.S2	25	25
MVKH.S2	25	25
ADDK.S2		16
SHL.S2	4	4
SPDP.S2	–	5
CMPEQSP.S2	–	2
MV.S2	1	1
LDW.D1	1	3
ADDAW.D1	–	2
STW.D1	–	1
ADDAD.D1	1	1
STW.D2	78	82
LDW.D2	93	93
ADDAW.D2	2	21
ADDAD.D2	2	2

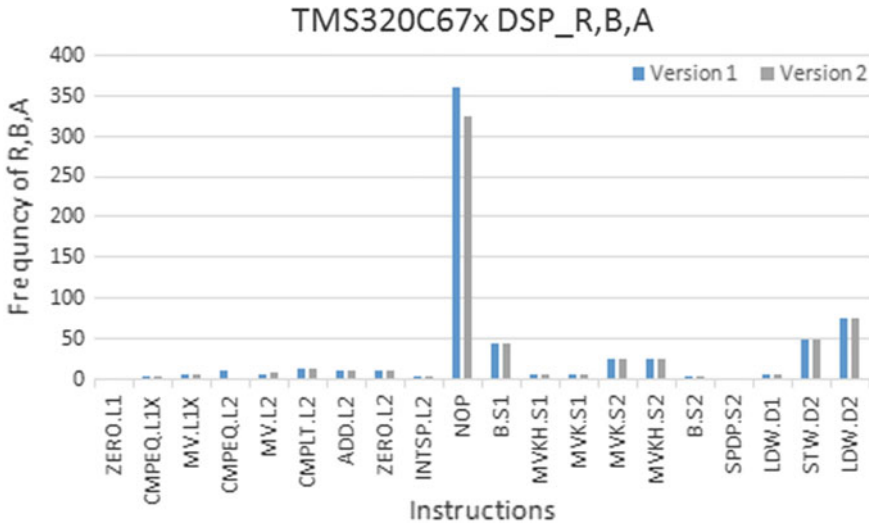


Fig. 1 Plot of frequency count versus assembly instructions generated for Reduct, Boundary, and Accuracy on DSP processor

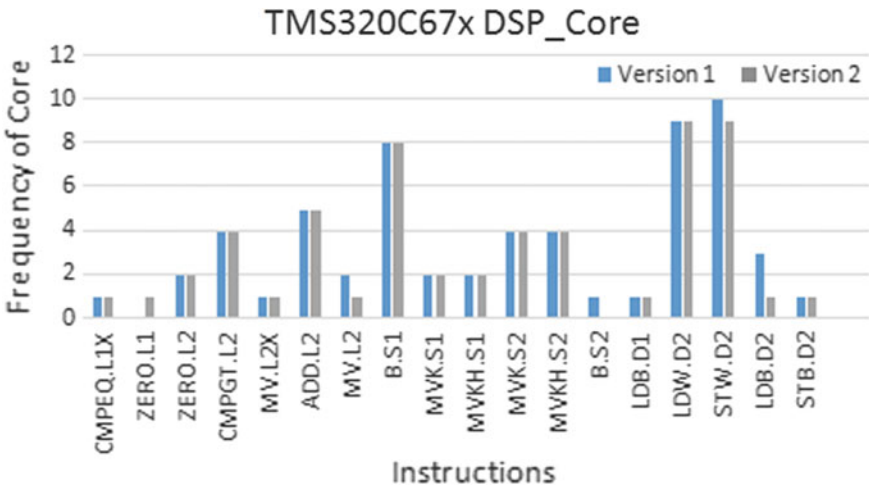


Fig. 2 Plot of frequency count versus assembly instructions generated for core on DSP processor

### 3.2 Discussion

The code profiling analysis of rough set theory suggests that the available platforms are not possible solutions for implementation of RST algorithm (for real-time application) due to numerous reasons.

**Table 4** Frequency count of assembly instructions generated for elementary set and crisp set on ARM cortex M4

Instruction	Version 1	Version 2
Push	3	3
Sub	3	3
Add.w	6	6
Add	20	30
Mov	17	22
Ldmia.w	4	4
Ldmia	4	4
Stmia	4	4
Stmia.w	4	4
Ldr	66	116
Movs	11	20
Str	5	22
Bl	2	17
Adds	12	12
Pop	2	2
b.n	6	6
Lsls	22	26
Lsrs	4	14
Cmp	8	9
Bne.n	3	4
Str.w	5	5
Ble.n	4	3
Bx	1	–
Beq.n	1	1
Nop	1	2

1. The DSP Platform is mainly used for digital signal processing applications. There is no provision for optimized calculation for rough set constructs that is set theory based specialized hardware. Similarly, in the case of ARM processor (basically used for embedded applications), there is no provision for the calculation of set elements. This is visible by several NOP instructions that are generated, wasting processor time, and power.
2. Rough set theory does not use Multiply and Accumulate (MAC) (.M unit) kind of calculations as shown in Table 6. So, possibility of silicon wastage can occur.
3. Lot of time is wasted in loading and storing of data. When the data size increases, it becomes a time-consuming process.

**Table 5** Frequency count of assembly instructions generated for  $L_A$  and  $U_A$  on ARM cortex M4

Instruction	Version 1	Version 2
Push	2	2
Add	5	34
Ldr	146	83
bl	14	25
Mov	35	27
Sub	24	1
Pop	2	1
bx	2	–
Word	16	–
str	47	14
b	10	–
lsl	31	–
cmp	24	20
bne	7	–
Beq	8	–
ble	4	–
Blt	5	–
Mov.w	–	1
Movs	–	31
Lsrs	–	15
Str.w	–	33
b.n	–	12
Lsls	–	31
Vldr	–	41
Ldr.w	–	78
Vmov	–	2
Vcvt.f32	–	2
Vcmp.f32	–	4
Bne.w	–	1
Beq.n	–	10
Adds	–	30
Vstr	–	1
Subs	–	13
Ble.w	–	4
Nop	–	3
Bne.n	–	4
Blt.n	–	5
Add.w	–	13
Ble.w	–	4

**Table 6** Classification of assembly instructions generated for elementary set and crisp set on ARM cortex M4

Sr. no.	Rough set theory constructs	.L	.M	.S	.D
1	Elementary and crisp set	ADD.L1X		MVK	LDW
		MV.L1X		MVKH	STW
		ZERO		SHL.S1X	ADDAD
		CMPGT		ADDK	ADDAW
		CMPEQ		SHL	
		ADD		MV	
		MV			
2	$L_A, U_A$	ZERO		B	LDW
		CMPGT		MVK	STW
		MV.L1X		MVKH	ADDAD
		ADD		ADDK	ADDAW
		MV		SHL.S1X	
		CMPEQ.L1X		SHL	
		XOR			
		MV.L2X			
		ADD.L2X			
CMPLT					
3	Reduct, accuracy, boundary	ZERO		B	LDW
		MV.L1X		MVK	STW
		CMPEQ.L1X		MVKH	ADDAW
		MV		SPDP	
		MV.L2X			
		ADD			
		CMPLT			
		CMPEQ			
		INTSP			
4	Core	ZERO		B	LDW
		CMPEQ.L1X		MVK	STW
		CMPGT		MVKH	LDB
		ADD			STB
		MV			
		MV.L2X			

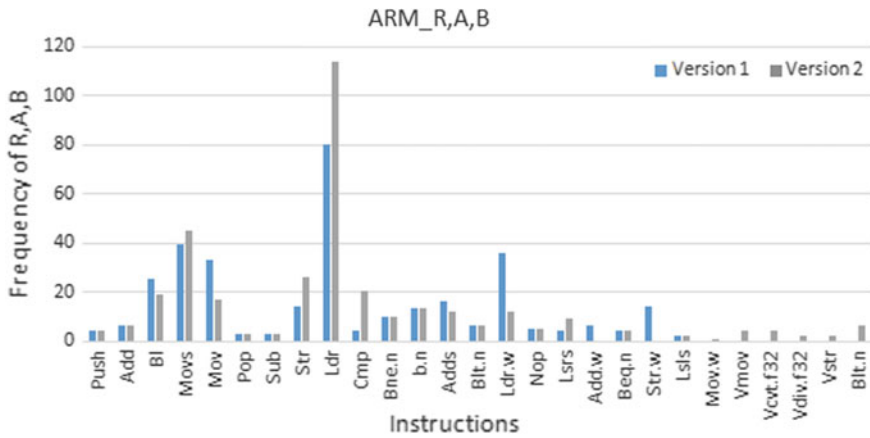


Fig. 3 Plot of frequency count versus assembly instructions generated for Reduct, Boundary, and Accuracy on ARM cortex M4 processor

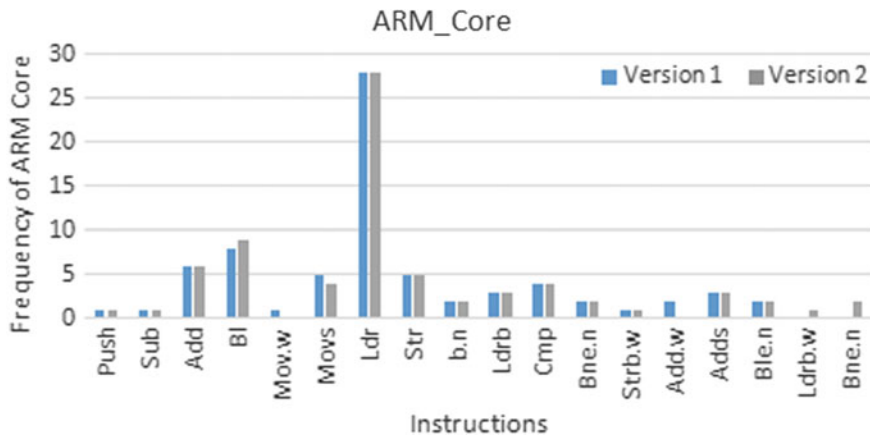


Fig. 4 Plot of frequency count versus assembly instructions generated for core on ARM cortex M4 processor

### 3.3 File Sizes

Table 7 shows the file sizes of code and output for various rough set technique constructs. This study suggests that output file generated is high when only six objects have been considered. As more and more data needs to be analyzed, the output file size will tremendously increase. This puts enormous computational load on existing hardware platforms due to requirements of extra memory. Therefore, the authors recommend a specific processor for analyzing inconsistent data for IoT with the support of specific instruction set for rough set theory.



**Table 7** Generated files and their sizes

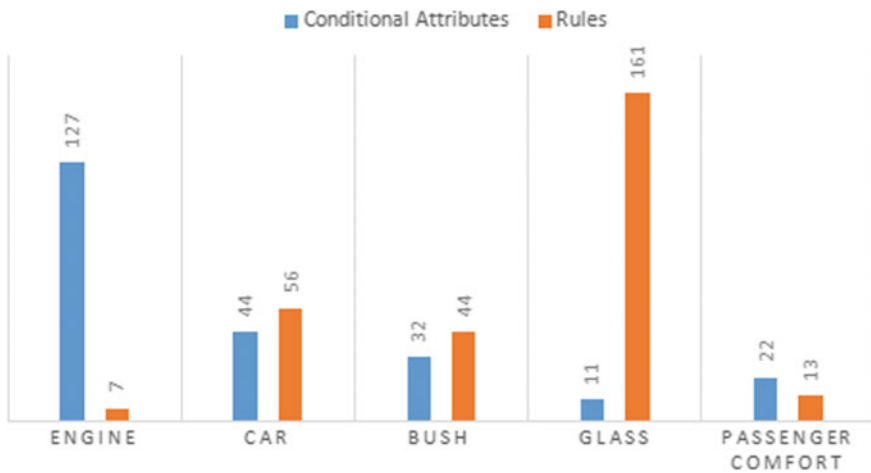
RST constructs	DSP C6713		Cortex M4	
	.C (KB)	.out (KB)	.C (KB)	.out (KB)
ES, CS	3	133	3	57
$L_A, U_A$	2	137	4	5
Reduct, Boundary, Accuracy	2	137	3	57
Core	2	129	2	57

## 4 Generated Rule Set for Data Sets From Different Applications of IoT

Using rough set technique, Table 8 tabulates the generated rule set for different data sets from the literature (standard examples from ROSE software [11]). Various examples of different number of conditional attributes have been taken. Figure 5 shows a

**Table 8** Generated rule set for data sets from different applications of IoT

Different data sets	Conditional attributes count	Decision attributes count	Decision attributes for which rules have been generated	Generated rules count
ENGINE	127	2	D1	7
CAR	44	6	DEC	56
BUSH	32	2	D	44
GLASS	11	7	Type of glass	161
PASSENGER COMFORT	22	1	D1	13



**Fig. 5** Plot of frequency versus assembly instructions generated for core on ARM cortex M4 processor

plot of frequency versus assembly instructions generated for Core on ARM Cortex M4 processor.

The number of rules generated depends on the elementary sets obtained. We have not shown the relationships here. It is available in standard textbook [15].

## 5 Conclusion

This study shows that there is no provision for optimized calculation for rough set theory on DSP and ARM processors. For the first time, the authors have proved by experimentation that there is a full justification for developing hardware coprocessor for supporting specific instruction set architecture for handling inconsistent information systems using rough set theory for IoT application. As edge based hardware is constrained in computational processing requirements, a support of specific instruction set can definitely suffice the power/energy requirement for dealing with inconsistent information system in IoT applications.

**Acknowledgements** The authors would also like to thank Mr. A. B. Patki, Ex-Senior Director/Scientist G and HoD, Ministry of Electronics and Information Technology, Government of India for his valuable suggestions and guidance. Authors also acknowledge the help and support of College of Engineering Pune for carrying out this work.

## References

1. Pawlak, Z. (1982). Rough sets. *International Journal of Computer and Information science*, 11, 341–356.
2. Pawlak, Z. (1984). Rough classifications. *International Journal of Man Machine studies* (No. 20)
3. Pawlak, Z. (1991). *Rough Sets: Theoretical aspects and reasoning about data*. Dordrecht: Kluwer Academic.
4. Jiye, Li, & Cercone, Nick. (2006). Assigning missing attribute values based on rough sets theory. In *Proceedings of IEEE International Conference on Granular Computing, GrC*, (Vol. 2006(May), pp. 10–12).
5. Riza, L. S., et al. (2014). Implementing algorithms of rough set theory and fuzzy rough set theory in the R package “RoughSets”. *Information Sciences*, 287, 68–89.
6. Hassan, Y. F. (2017). Deep learning architecture using rough sets and rough neural networks. *Kybernetes*, 46(4), 693–705. <https://doi.org/10.1108/K-09-2016-0228>.
7. Zhang, Q., Xie, Q., & Wang, G. (2016). A survey on rough set theory and its applications. *CAAI Transactions on Intelligence Technology*, 1, 323–333.
8. Bello, R., & Falcon, R. (2017). “Rough Sets in machine learning: A review”, chapter in studies in computational. *Intelligence*. [https://doi.org/10.1007/978-3-319-54966-8\\_5](https://doi.org/10.1007/978-3-319-54966-8_5).
9. Jiang, H. Study on the application of rough sets theory in machine learning. In *Proceedings of Second International Symposium on Intelligent Information Technology Application*. <https://doi.org/10.1109/IITA.2008.154>
10. Mahajan, P., Kandwal, R., & Vijay, R. (2012). Rough set approach in machine learning: a review. *International Journal of Computer Applications*, (0975-8887) 56(10)

11. ROSE 2 User guide. (2017). Retrieved June 25, 2017, from <http://idss.cs.put.poznan.pl/site/fileadmin/projects-images/rosemanual.pdf>.
12. Abbas, Z., & Burney, A. (2016). A survey of software packages used for rough set analysis. *Journal of Computer and Communications, 4*, 10–18.
13. Agarwal, V., Patil, R. A., Patki, A. B. Architectural considerations for next generation iot processors. *Accepted for Publication in IEEE Systems Journal*. <https://doi.org/10.1109/JSYST.2018.2890571>
14. Jyoti, A. (2017). *Code profiling for RST algorithm on DSP and embedded processors*. M.Eng thesis, College of Engineering Pune, India.
15. Munakata, T. (2008). Rough sets. In *Springer: Fundamentals of the new artificial intelligence neural, evolutionary, Fuzzy and More* (2nd ed., pp. 162–202).

# Design and Analysis of IoT-Based System for Crowd Density Estimation Techniques



Ajitesh Kumar and Mona Kumari

**Abstract** In this paper, we present an IoT-based solution that can reduce the complexity of crowd estimation. About the human crowd estimation many techniques are in existence but nowadays more work is going on in this field because this is era of IoT and most of the organization is shifted toward IoT-based system. So in our proposed system we are using the Raspberry Pi-3 which are having quad-core processor that can be very useful and gives better result and accurate number even when the humans are very close to each other. This IoT-based model can easily be implemented in crowded areas and monitor the same. The camera module in this model also helps to differentiate between human and other bodies. As this is a mobile model, it can be easily fixed on the walls of street light and in the time of darkness or in night the camera captures clear images for process in the presence of street light. So that this model gives better result almost 70% better result in compare to exiting approaches.

**Keywords** Zigbee · Crowd density · Raspberry Pi-3 · IoTBCET · RFID

## 1 Introduction

Our objective of this work to reduce the complexity of crowd estimation by using IoT-based system with Raspberry Pi that can easily count humans. The IoT-based localization is a process of counting the humans by their position and movement within a network by using mathematical techniques [1, 2]. The system is able to perform by location sensing, using RFID or target tracking, and sometimes both [3, 4]. Crowd counting and monitoring are very useful to avoid the accident. This device handling technique plays a very important role in estimating crowd and gives a very good result in when compared to the existing approaches. In this approach,

---

A. Kumar (✉) · M. Kumari  
GLA University, 17 Km Stone, NH-2, Mathura-Delhi Road, Mathura 281406, UP, India  
e-mail: [ajitesh.kumar@gla.ac.in](mailto:ajitesh.kumar@gla.ac.in)

M. Kumari  
e-mail: [mona.kumari@gla.ac.in](mailto:mona.kumari@gla.ac.in)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_29](https://doi.org/10.1007/978-981-15-0694-9_29)

e.g. Counts of people is used for crowd control.



**Fig. 1** Shows the crowded area

everyone have mobile devices and our system received signal between receiver and sender which are very useful data for our approach. So that the objective of this paper is to be very clear that we develop a system and discuss the significant factors affecting the RFID identification (Fig. 1).

Once the effectiveness of human crowd is understood then we easily get the information from all users.

The system can estimate in real time and based on RFID, (IoT based) the previous proposed methods cannot count the number of people and track the crowds in real time. If the number of individuals increases, the system degrades drastically [5].

A large group of people is called crowd, who are available in a particular area. In general, places like airport, daily market, bus stations, railway station are very necessary and it is a difficult task to identify the unwanted person over thousands of people. It is very difficult for human to count the head and identify over thousands of gathering manually [4].

## 2 Literature Survey and Research Gap

In smart-system-based counting, people are usually avoided in all aspects and they are not agreed to share the personal location in the system which is main challenge for head counting. Most of the systems used the data which are given by people in crowd and are not guaranteed to share so that we gives the some incentive or some offers to the crowd so that they can share the information which are very necessary to head counting [6, 7].

### Discussion on Some Similar Technique:

A Wi-Fi based where they allow crowd to play a geographical game and based on that they collect the information from the users. They allow playing only in Wi-Fi enabled area so that crowd may be bound and also it will be a challenge for that [8].

**Table 1** Shows comparative study of different techniques

Sensing facilities	Related work/platform	Main features
Participatory	WISP-based [9] -RFID	Provides framework on crowd based system, provide geographical data on mass event gathering
	Hand phone crowd monitoring	Provide collaborative Wi-Fi and Bluetooth system
Nonparticipatory	Electronic Frog Eye-Wi-Fi [4]	Utilizes channel-based state information to estimate crowd density
	Wi-Counter [11] -Wi-Fi	Provide three-phase iterative

Related works on the field RFID-based system are discussed below in the table (Table 1):

Crowd dynamics for analysis is also very complex topics nowadays [9]. In this paper, an effective technique is used and gives better results over DOE.

In the DOE they used the crowd dynamic factors that can reduce the overall complexity. So this technique is useful in nondance areas. The Zigbee chipset used in the model is dramatically changed on the result as discussed and shown by others in their papers [9, 10]. In the work of RF-based H-CDE is shown in the table they said in their paper that RF-tagged devices are used but the major challenge in this regard is to be difficulties of tagging the RF tag in the crowded areas [11]. Maybe the person is not interested to involve or participate in this model so it is necessary to ask everyone about the benefits of this model [12].

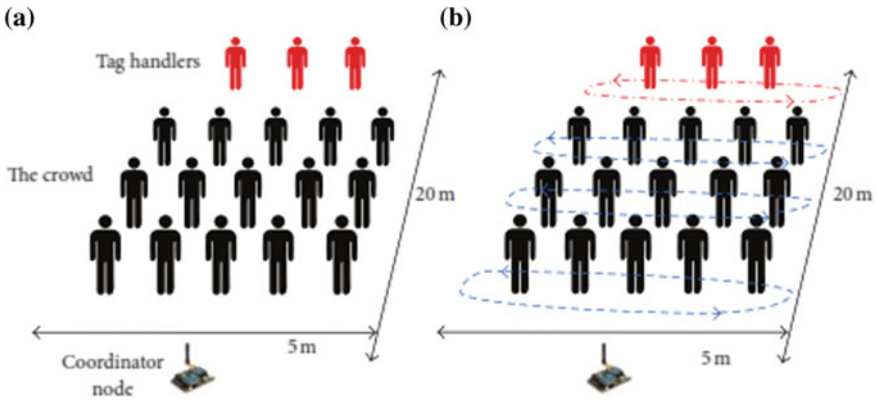
The visual sensors that have been widely used are wireless sensor network, computer vision, smart camera, sensor fusion and few more; and the nonvisual sensors are Call Data Records, Wi-Fi Signals Measurement, Smart Eva track, Social Network, and Bluetooth, etc. Automatic crowd understanding has a massive impact on several applications including surveillance and security, situation awareness, crowd management, public space design, intelligent, and virtual environments [13, 4].

The motivation behind this approach is non-accurate values in the existing approaches and we are going to use Raspberry Pi in the place of Arduino Uno because Wi-Fi facilities are available with Raspberry Pi.

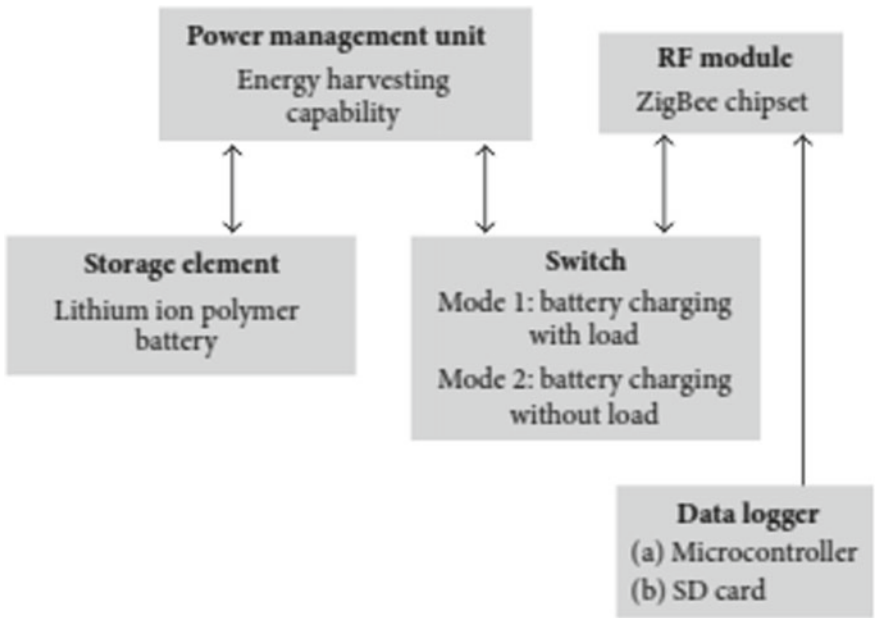
### 3 Proposed Model and Experimental Framework

The proposed system uses the Raspberry Pi-3 model B RASP-3 motherboard with the Wi-Fi facilities for faster process. This motherboard has four USB ports and one HDMI port is built with the 64-bit processor. It is like a quad-core CPU with micro SD slots. It takes less power for operation and easily build up the process.

RFID tags are very useful and can be easily tagged on the items. And nowadays every mobile system have Wi-Fi that is useful for identification of movement of human in the crowd (Figs. 2 and 3).



**Fig. 2** Figure a shows the experimental setup where all elements are static in nature and in figure b and c the crowd having movements within a given area



**Fig. 3** Block diagram of the system

This is a block diagram of the system where we use batteries which are charged by harvesting (Fig. 4).

In this system, we also using the camera-based image identifier so that we can easily differentiate between human and other bodies like robot, etc. (Fig. 5).

That shows the dense population in one place that creates a problem on camera-based system that cannot measure the exact image so that we need IoT-based system that are used the human system RFID tags. This shows how we differentiate the human body and other items. This is very useful technique for identifying such type of process (Fig. 6).

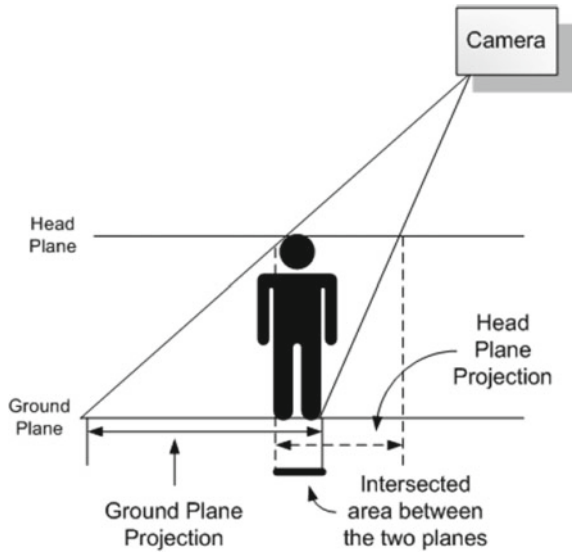


Fig. 4 Shows camera module

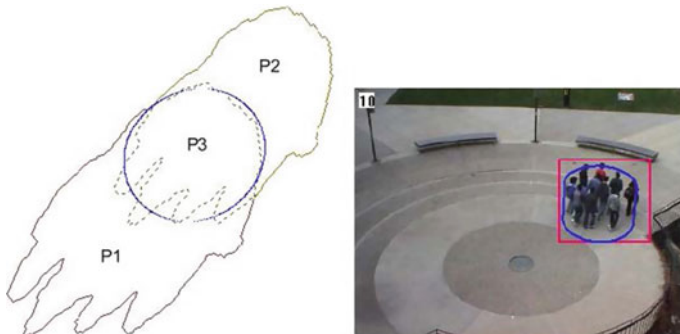
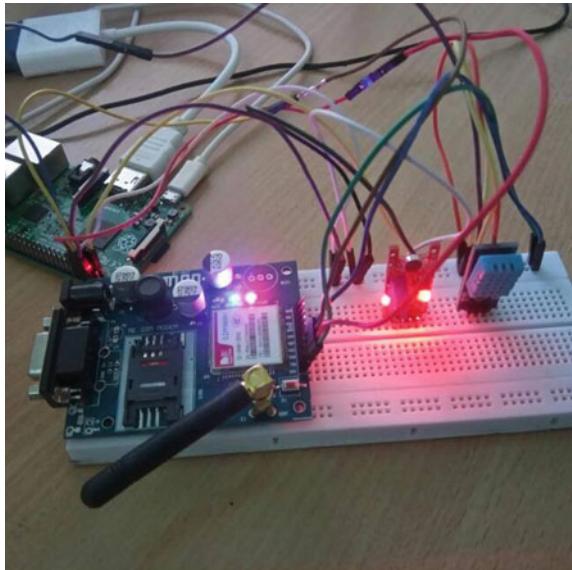


Fig. 5 Shows dance areas



**Fig. 6** Shows working module



**Table 2** Shows level and factors with count level

Factors	Level 1	Level 2	Level 3
Crowd size (Human)	5	10	15
Crowd pattern	Scattered	Lumped	-
Location (m)	10	20	30
Number of tags	1	2	3

### 4 Results & Discussion

For better understanding of this proposed model we need some Experimental setup (Table 2).

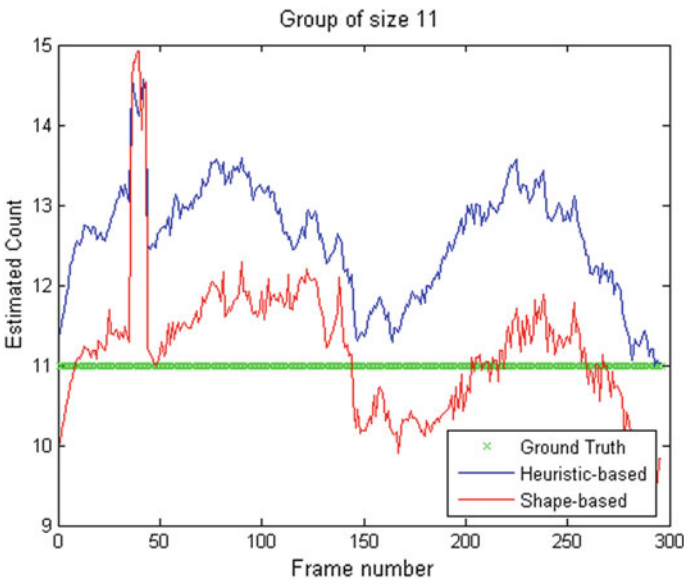
We observe three different scenes of the crowd and seen eight different positions. The camera height was varied from 29 feet to 80 feet and the tilt of camera was varied from 30 to 40° and the most crowded scene is up to 50 people (Table 3, Figs. 7 and 8).

### 5 Conclusions

So in this proposed system we can count crowd of people accurately in real time. In this proposed model, we are easily counting the humans in dance areas by using RFID tags and the camera module can easily identify the human body. There are some occasional problems with existing method that can be resolved in this approach. The

**Table 3** Shows number of frames with shape error per frame

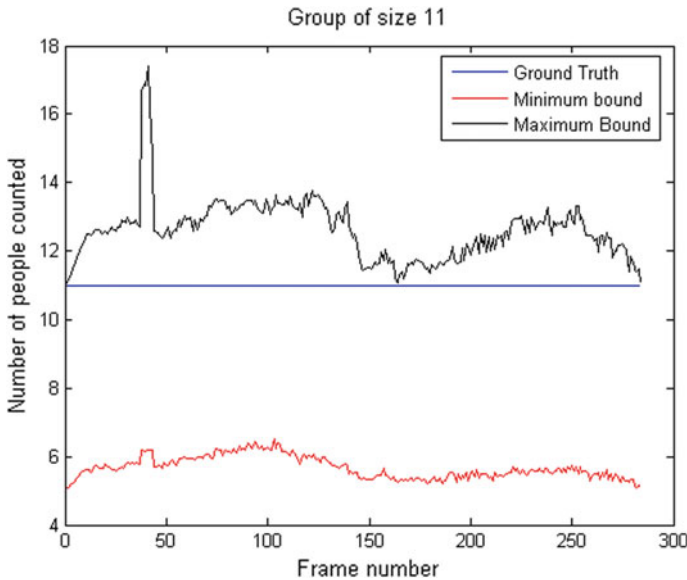
Group size (No. of peoples)	No. of frames	Heuristic error per frame	Shape error per frame
8	332	1.44	1.17
9	530	1.51	1.30
8	372	1.04	0.85
11	354	1.81	0.72
9	384	0.71	0.83
10	156	1.53	1.24
10	224	1.86	1.03



**Fig. 7** Plotting of exact counts over a group of 11 people

experimental result shows that proposed algorithm has better result in comparison with existing one.

As future aspects, we can work upon the data analytics concepts where we can test with more object images, and we also work upon the movable devices based on IoT that can move if system requires.



**Fig. 8** Plotting of exact counts over a group of 11 people in dance areas

## References

1. Boukerche, A., Oliveira, H. A. B. F., Nakamura, E. F., & Loureiro, A. A. F. (2007). Localization systems for wireless sensor networks. *IEEE Wireless Communications*, 14(6), 6–12.
2. Ni, L. M., Liu, Y., Lau, Y. C., & Patil, A. P. (2003). LANDMARC: Indoor location sensing using active RFID. In *Proceedings of the 1st IEEE International Conference on Pervasive Computing and Communications (PerCom'03)* (pp. 407–415). Fort Worth, Tx, USA: IEEE. Retrieved March 2003.
3. Dian, Z., & Ni, L. M. (2009). Dynamic clustering for tracking multiple transceiver-free objects. In *Proceedings of the IEEE International Conference on Pervasive Computing and Communications (PerCom'09)*, Galveston, Tx, USA (pp. 1–8). Retrieved March 2009.
4. Arai, M., Kawamura, H., & Suzuki, K. (2010). Estimation of ZigBee'sRSSI fluctuated by crowd behavior in indoor space. In *Proceedings of the SICE Annual Conference (SICE'10)*, Taipei, Taiwan (pp. 696–701). Retrieved August 2010.
5. Xu, C., Firner, B., & Moore, R. S., et al. (2013). SCPL: Indoor device free multi-subject counting and localization using radio signal strength. In: *Proceedings of the 12th International Conference on Information Processing in Sensor Networks (IPSN'13)*, Philadelphia, Pa, USA (pp. 79–90). Retrieved April 2013.
6. Huang, C.-N., & Chan, C.-T. (2011). ZigBee-based indoor location system by k-nearest neighbor algorithm with weighted RSSI. *Procedia Computer Science*, 5, 58–65.
7. Bahl, P., & Padmanabhan, V. N. (2000). RADAR: An in-building RF based user location and tracking system. In *Proceedings of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'00)* (Vol. 2, pp. 775–784). Tel Aviv, Israel: IEEE. Retrieved March 2000.
8. Oka, A., & Lampe, L. (2010). Distributed target tracking using signal strength measurements by a wireless sensor network. *IEEE Journal on Selected Areas in Communications*, 28(7), 1006–1015.

9. Liu, X.-L., Chen, Y.-G., Jing, X.-R., & Chen, Y.-W. (2010). Design of experiment method for microsatellite system simulation and optimization. In *Proceedings of the International Conference on Computational and Information Sciences (ICIS'10)* (pp. 1200–1203), Chengdu, China: IEEE. Retrieved December 2010.
10. Litvinski, O., & Gherbi, A. (2013). Open stack scheduler evaluation using design of experiment approach. In *Proceedings of the 16<sup>th</sup> IEEE International Symposium on Object/Component/Service Oriented Real-Time Distributed Computing (ISORC'13)*, Paderborn, Germany (pp. 1–7). Retrieved June 2013.
11. Fadhlullah, S. Y., & Ismail, W. (2015). Solar energy harvesting design framework for 3.3 V small and low-powered devices in wireless sensor network. In *Proceedings of the 1st International Conference on Telemetric and Future Generation Networks* (pp. 89–94). Kuala Lumpur, Malaysia: IEEE. Retrieved May 2015.
12. Curtis, S., Guy, S. J. Zafar, B., & Manocha, D. VirtualTawaf: A case study in simulating the behavior of dense, heterogeneous crowds. In *Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCV'11)* (pp. 128–135). IEEE.
13. Karamouzas, I., Skinner, B., & Guy, S. J. (2014). Universal power law governing pedestrian interactions. *Physical Review Letters*, 113(23), Article ID 238701.

# Video-Transmission-Based Condition Monitoring of Solar Panels Using QR Code



Akash Singh Chaudhary, Isha and D. K. Chaturvedi

**Abstract** Sun is a clean and renewable source of energy and produces solar energy by converting solar radiations into useful electricity through solar cells. Solar energy is obtained from solar radiations by the phenomenon of photoelectric effect. Solar panels are installed in open atmospheric conditions and undergo with environmental effects. These solar panels are subjected to various defects and faults during operation; therefore, proper condition monitoring is needed. Data loggers are used to remotely monitor the condition of solar panels, i.e., voltage, current, temperature, and other atmospheric parameters on the screen of personal computer. QR codes are normally used in commercial applications for information exchange but in this research work a novel technique based on QR code is used to view the variations of values and graphs for different parameters of solar panels via well-designed recording system through an Android app generated for particular QR code. By scanning the QR code, live variations of graphs and values in video form for solar panel data can be visualized which are not possible after time of visualization through data logger is lost. Therefore the research work presents a unique technique for visualizing recorded data in video form for solar panels as off-line whenever needed through a simple Android mobile generated QR code.

**Keywords** Condition monitoring · Data logger · Solar panel · QR code · Video transmission

---

A. S. Chaudhary (✉) · Isha · D. K. Chaturvedi  
Department of Electrical Engineering, Faculty of Engineering, Dayalbagh Educational Institute,  
Agra, India  
e-mail: [akashsinghchaudhary@gmail.com](mailto:akashsinghchaudhary@gmail.com)

Isha  
e-mail: [ishasingh268@gmail.com](mailto:ishasingh268@gmail.com)

D. K. Chaturvedi  
e-mail: [dkc.foe@gmail.com](mailto:dkc.foe@gmail.com)

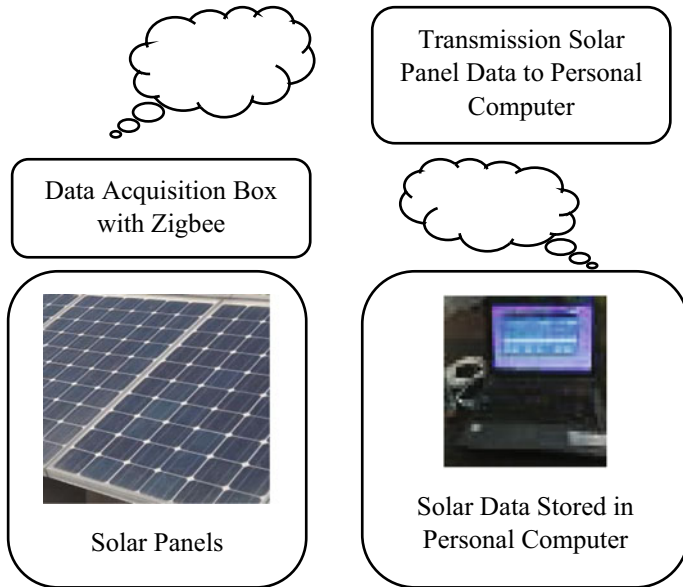
© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_30](https://doi.org/10.1007/978-981-15-0694-9_30)

## 1 Introduction

Energy obtained by the sun has many advantages compared to conventional sources and is available in abundance. Solar cells absorb solar radiations emitted by sun and produce solar energy [1, 2]. A solar photovoltaic module consists of solar cells which are connected in a well-arranged manner of series and parallel combination called solar module. The electrical energy output obtained from these solar panels is used to charge batteries through charge controller and then fed to load through inverter [3, 4]. During their operation, solar panels are exposed to large change in their operating temperature, voltage, current, and output power. The overall effect of these environmental conditions is reflected in terms of their performance and operating efficiency [5, 6]. The efficiency of solar panels is affected by shading effect of tree, building or tower [7, 8], degradations, and aging effect in solar photovoltaic components with reduction in overall power output of complete solar photovoltaic system. To operate solar panels in a satisfactory region the condition of solar panels need to be monitored [9–11]. The output voltage, output current, output power, atmospheric condition, solar irradiance, and other atmospheric parameters can be remotely monitored through data logging system. Data loggers provide an efficient online tool for monitoring the live variations of different parameter values of solar panels showing the corresponding change in their values. These data loggers are also useful in monitoring the condition of solar batteries remotely with the voltage, current, state of charging and discharging [12, 13], and in personal computer [14]. These parameter values are stored in MS Excel file format (.csv) in a notepad file in the personal computer through which data logging is performed with an advantage to view only the stored values in future. The personal computers receive the information sent from the solar panel data acquisition box through transmission devices such as Zigbee and GSM [15, 16]. Data logger use Zigbee and GSM for condition monitoring system [17]. The static graphs can be generated from these values in MS Excel file from the stored solar data [18]. This transmission of solar panel data has limitations time delay and data loss in the transmission of recorded data in video form [19]. QR code has numerous applications in transmission and video sharing with reduced complexity in data transmission [20].

## 2 Concept of Condition Monitoring Through Data Logger and Data Transmission Using QR Code

Solar photovoltaic generation system can be classified into two main categories they are off-grid (standalone) solar photovoltaic generation system and grid-connected (online) solar photovoltaic generation system. The main components of an off-grid solar photovoltaic generation system are Solar panels, Solar charge controller, Solar battery, and Solar inverter. The grid-connected solar photovoltaic generation system is directly connected to load with no solar batteries [21, 22]. Data loggers use wireless techniques for remote condition monitoring of solar panel on the screen of a personal



**Fig. 1** Condition monitoring of solar panels through data logger

computer [23, 24]. The advantages of ZIGBEE and GSM are more as compared to other wireless transmission devices so they are frequently used nowadays [25, 26]. The data received by data logger is stored in personal computer and can be retrieved later but the live variations of values and graphs of output voltage, output current, output power, atmospheric and panel operating temperatures, solar irradiance, and other parameters cannot be visualized online [27]. Each QR code has several applications in almost every field due their simplicity, easy of transfer, and sharing. Some of the advantages of QR code are fast scanning, smaller in size, more capacity to store data, support to many languages (numeric, alphanumeric, kanji etc.), and accessible through 360° [28]. The QR can store maximum 7089 maximum characters, with 4296 alphanumeric characters, with 2953 binary bytes and 1817 Kanji characters. The additional advantages of a QR code are that it can store texts, contacts, URL, shareable link, email, etc. so used in educational institutions, business, medical, and security [29]. QR code is a two-dimensional barcode having more storage capacity than one-dimensional barcode because in QR code the data can be stored in both direction (horizontal and vertical). There are two regions, namely, function pattern region and encoding region in QR code. The function pattern region contains finder, separator, timing patterns, and alignment patterns while the information is stored in encoding section of a QR code. By providing a shareable link to QR code stored information in it can be shared [30, 31]. These QR codes can be scanned through a QR code scanner through Android mobile phones to obtain the stored data [32]. The following block diagram shown in Fig. 1 represents the complete mechanism of remote condition monitoring of solar panels.

### 3 Video Transmission and Generation of QR Code for Solar Panels

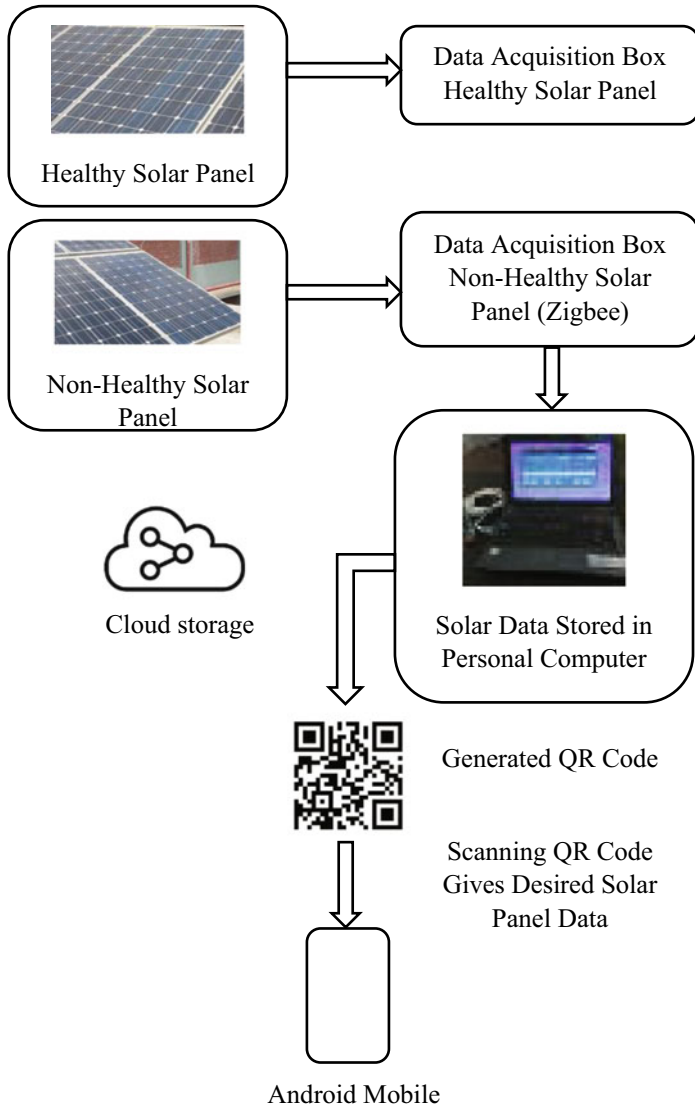
The work for the experiment is performed on a personal computer connected to the solar panels installed at the roof of Electrical Engineering Library, Faculty of engineering, Dayalbagh Educational Institute, Agra. Two solar photovoltaic panels out of complete solar photovoltaic system are selected. These solar panels are installed in open atmosphere and are subjected to various environmental conditions. One panel has no shading effect, no dust, and no defect so it is in healthy condition (Healthy Panel). The other panel faces shade of tree leaves, has bird dropping, cement depositions, and hotspots so it is in unhealthy condition (Non-Healthy Panel). The remote condition of these two panels, i.e., Healthy Panel and Non-Healthy Panel is performed using a data logger. The communication of solar panels data (variations in output current, output voltage, and output power with atmospheric and panel operating temperature) is done through a Zigbee. The solar panel data is monitored and stored in the personal computer connected to Internet. These variations and graphs can be monitored online and automatically stored in MS Excel sheet (.csv format) in the personal computer. The research work is done for recording the online solar panel data in video form through Google Chrome and generating a QR code with the desired recorded video file of solar panel data. The following Table 1 shows name plate rating of solar module used in the experiment.

Figure 2 shows complete block diagram of the algorithm used in the experimental work. The solar panel data for both healthy as well as non-healthy panels are acquired and stored in.CSV format in a Notepad file in personal computer through Zigbee.

**Table 1** Name plate rating of solar panels used in experiment

S. no.	Name	Value	
1.	Company	Bhel, India	
2.	Module No.	L20220	
3.	P <sub>maximum</sub>	220 W Power	
4.	V <sub>maximum power</sub>	29 V	
5.	I <sub>maximum power</sub>	7.6 A	
6.	V <sub>open circuit</sub>	36 V	
7.	I <sub>short circuit</sub>	8.3 A	
8.	Maximum system voltage	1000 V	
9.	Bypass diode rating (current)	15 A	
10.	Maximum series fuse current protection	15 A	
11.	Solar Module STC (Standard test conditions):		
	A.	Insolation	1000 W/m <sup>2</sup>
	B.	AM	1.5 spectrum
	C.	Cell Temp.	25 °C





**Fig. 2** Block diagram of algorithm used for condition monitoring and video transmission of solar panel using QR code



**Fig. 3** Different parameters value window, notepad file window, and graph window as displayed by data logger for healthy and non-healthy solar panel, respectively

Data logger window shows the online condition monitoring with online variations of different parameters in graphical form for healthy and non-healthy solar panels. Now video recording for these online condition monitoring windows are obtained with help of Google Drive as cloud storage and a QR code is generated by providing the shareable link of the Google cloud storage. In the next step, an Android app is developed which is linked with the generated QR code to access the stored video transmitted data file of solar panels.

### ***3.1 Video Recording of Solar Panel Data in Personal Computer***

The solar panel data for healthy and non-healthy panels is obtained through data logger with Zigbee in the personal computer having Internet connection. The video recording of the variations in values and graphs for different parameters of both solar panels which are obtained by data logger in personal computer are recorded by Google Chrome extension (screencastify). After completing the recording, the recorded video is saved in Google Drive by manual operations on personal computer. Figure 3 shows snapshots of recorded value window on-screen of personal computer, Notepad file stored in memory of personal computer and graphs of different parameters for solar panels as displayed on personal computer screen through data logger.

### ***3.2 Generation of QR Code for Recorded Video File***

After saving the video recorded file in Google Drive the shareable link is generated and QR code is generated which is copyrighted for security reasons. The generated QR code is attached to an Android web application for accessing the recorded video file by simple scan and registration process. The web application used for generation

**Fig. 4** QR code generated for remote condition monitoring of solar panels



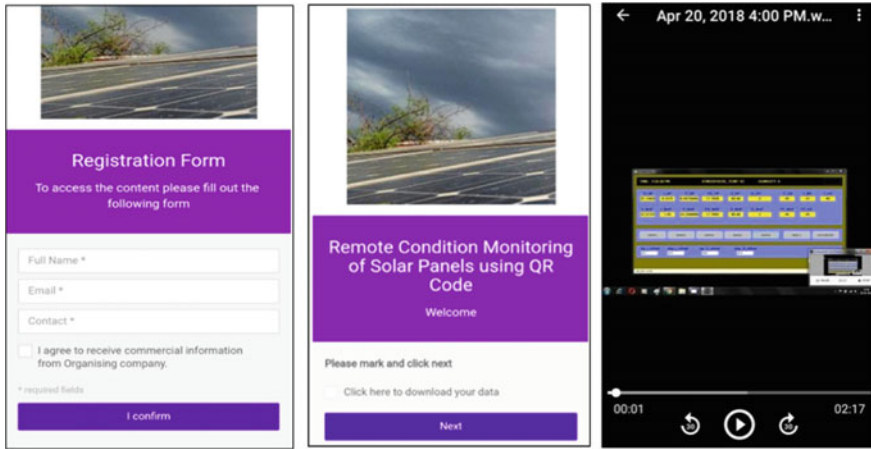
and accessing recorded video file through scanning is Unitag. The web application registers user whenever QR code is scanned and provide a user friendly platform for remote monitoring of stored video data of solar panels for better visualization. The live variations of values and graphs for solar panels can be viewed by scanning the following generated QR code shown in Fig. 4.

### ***3.3 Implementation of Condition Monitoring and Video Transmission System for Solar Panels Using QR Code***

The generated QR code is scanned by the designed web application through Android mobile phone have QR code scanner. The Fig. 5 shows the snapshots after scanning the QR code from Android mobile. In Fig. 5 first snapshot shows registration form after scanning QR code from Android mobile second snapshot shows check mark and the last snapshot shows that recorded video file is download and ready to play.

## **4 Results and Discussion**

The main objective is to design an off-line condition monitoring system for solar panels to view the real-time variations of values and graphs that were shown by data logger on the screen of personal computer during real-time condition monitoring. The technique used in the research work is useful in remotely monitor the condition of solar panels by user friendly QR code. The live variations of output current, output voltage, output power, solar irradiance, operating temperature of solar panels, and atmospheric temperature with their corresponding graphs are stored in video format in a generated QR code. This QR code can be transmitted and shared easily through any communication medium. Moreover, one of the advantages of this technique is



**Fig. 5** Snapshots of registration form, check mark to download data and downloaded window for solar panels after scanning QR code from android mobile

**Table 2** Details of recorded video file for solar panel data and generated QR code

S. no.	Size	Format	Resolution
Recorded video file	5.62 MB	(.webm)	1280 × 720 pixels with 30 fps
Generated QR code	53 KB	(.png)	300 × 300 pixels

reduction in size of the content, i.e., the generated QR code need less storage capacity compared to the recorded video format data. Table 2 shows detail of recorded video file before and after conversion into QR code.

## 5 Conclusions

The main objective of the research work is to design a QR code which is used to visualize the live variations of solar panel data in the form of a recorded video data file by scanning it through a QR code scanner with an Android mobile. During experimental work it has been found and concluded that space required to store a video in QR code is less as compared to normal format because video file is stored in image format in QR code, whereas normally video file has video format. The findings have been shown in Table 2 above in results and discussion section. It represents an efficient and novel technique to transfer the solar panel data through QR code and experimental work can be extended to remotely monitor the condition of solar panels for a particular duration. This work used free services of Internet web application and QR code generation websites; therefore, has limitation of recording

duration of solar panel data. One major advantage of this technique is to see the live variations of values and graphs in the same display format as data logger is showing on the screen of personal computer. Therefore, it is an off-line visualization method to see the online recorded data by scanning a QR code through Android mobile phone anywhere and at any time. The research work has been carried out with the help of free service available from websites such as Google Chrome and QR code generator. These websites have some limitation for free users so if the algorithm used in the research work is applied on professional means then a new condition monitoring strategy may arise in the area of photovoltaic. Main advantage of such new technique is off-line visualization of real-time variations on values and graphs which can be stored and easily accessed using generated QR code. Less space requirement is another advantage of storing solar panel information in QR code.

## References

1. Redfield, D. (1978). Solar energy and conversion. *Technology and Society, IEEE Journals and Magazines*, 6(23), 4–9.
2. Biran, D., & Braunstein, A. (1976). Solar radiation and energy measurements. *IEEE Transactions on Power Apparatus and Systems*, 95(3), 791–794.
3. Bouraiou, A., Hamouda, M., & Chaker, A., et al. (2015). Modeling and simulation of photovoltaic module and array based on one and two diode model using matlab/simulink. In *The International Conference on Technologies and Materials for Renewable Energy, Environment and Sustainability* (Vol. 74, pp. 864–877). Elsevier, Energy Procedia.
4. Machado Neto, L. D. B., Cabral, C. V. T., Oliveira Filho, D., et al. (2004). Monitoring of photovoltaic systems for performance evaluation and fault identification. In *2004 IEEE/PES Transmission & Distribution Conference & Exposition: Latin America, Sao Paulo, Brazil* (pp. 360–365).
5. Ali, M. H., Rabhi, A., Hajjaji, A. E., et al. (2017). Real time fault detection in photovoltaic systems. In *8th International Conference Sustainability in Energy and Buildings, SEB-16, Turin, Italy* (Vol. 111, pp. 914–923). Energy Procedia, Elsevier.
6. Dubey, S., Sarvaiya, J. N., & Seshadri, B. (2013). Temperature dependent photovoltaic (PV) efficiency and its effect on PV production in the world a review. *PV Asia Pacific Conference, Energy Procedia, Elsevier*, 32, 311–321.
7. Torres, J. P. N., Nashih, S. K., Fernandes, C. A. F., et al. (2016). The effect of shading on photovoltaic solar panels. *The Journal Energy Systems*, 1–14. Springer.
8. Chaudhary, A. S., & Chaturvedi, D. K. (2017). Observing hotspots and power loss in solar photovoltaic array under shading effects using thermal imaging camera. *International Journal of Electrical Machines and Drives*, 3(1), 15–23.
9. Chaturvedi, D. K., & Chaudhary, A. S. (2017). Condition monitoring of solar photovoltaic panels using infrared thermography. *Smart Energy, Genesis Info-Media*, 4(3), 34–36.
10. Dhoke, A., Sharma, R., & Saha, T.K. (2016). Condition monitoring of a large-scale PV power plant in Australia. In *IEEE Conference Power and Energy Society General Meeting, (PESGM), Boston, MA, USA*.
11. Kim, H. J., Lee, J. H., Baek, D. H., et al. (2017). A study on thermal performance of batteries using thermal imaging and infrared radiation. *Journal of Industrial and Engineering Chemistry, Elsevier*, 45(25), 360–365.
12. Anwari, M., Dom, M. M., & Rashid, M. I. M. (2011). Small scale PV monitoring system software design. *ICSGCE Chengdu, China, Energy Procedia, Elsevier*, 12, 586–592.

13. Petrescu, C., Lupu, C., Tudor, F.S., et al. (2013). Data acquisition system for recording of photovoltaic panel power. In *2nd IEEE, International Conference on Systems and Computer Science (ICSCS), Villeneuve d'Ascq, France* (pp. 26–27).
14. Purwadi, A., Haroen, Y., & Ali, F. Y. (2011). *Prototype development of a low cost data logger for PV based LED street lighting system* (pp. 1–5). Bandung, Indonesia: IEEE International Conference on Electrical Engineering and Informatics.
15. Katsioulis, V., Karapidakis, E., Hadjinicolaou, M., et al. (2011). Wireless monitoring and remote control of PV systems based on the ZigBee protocol. In *International Federation for Information Processing (IFIP) Advances in Information and Communication Technology* (Vol. 349, pp. 297–304). Springer.
16. Shariff, F., Rahim, N. A., & Ping, H. W. (2015). Zigbee-Based data acquisition system for online monitoring of grid-connected photovoltaic system. *Journal of Expert Systems with Applications, Elsevier*, 42, 1730–1742.
17. Asif, M., Ali, M., Ahmad, N., et al. (2016). Design and development of a data logger based on IEEE 802.15.4/ZigBee and GSM. In *Proceedings of the Pakistan Academy of Sciences, A. Physical and Computational Sciences* (Vol. 53, No. 1, pp. 37–48).
18. Sehgal, V. K., Nitin, & Chauhan, D. S. (2009). Smart wireless temperature data logger using IEEE 802.15.4/ZigBee Protocol. In *2008 IEEE Region 10 Conference, TENCN, Hyderabad, India*.
19. Kazemian, H. B. (2009). An intelligent video streaming technique in ZigBee wireless. In *IEEE International Conference on Fuzzy Systems, FUZZ-IEEE, Jeju Island, South Korea* (pp. 121–126).
20. Lien, S. F., Wang, C. C., Su, J. P., et al. (2014). Logistical remote association repair framework using smartphones based on the android platform. In *IEEE 2014 Int. Symposium Computer, Consumer and Control, Taichung, Taiwan* (pp. 1191–1194).
21. Rajeev, A., & Sundar, K.S. (2013). Design of an off-grid PV system for the rural community. In *IEEE 2013 International Conference on Emerging Trends in Communication, Control, Signal Processing & Computing Applications (C2SPCA), Bangalore, India*.
22. Mohammed, A. Y., Mohammed, F. I., & Ibrahim, M. Y. (2017). Grid connected photovoltaic system. In *IEEE, International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE), Khartoum, Sudan*.
23. Naeem, M., Anani, N., Ponciano, J., et al. (2011). Remote condition monitoring of a PV system using an embedded web server. In *2nd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies, Manchester, UK* (pp. 1–4).
24. Stauffer, Y., Ferrario, D., Onillon, E. (2015). Power monitoring based photovoltaic installation fault detection. In *4th IEEE International Conference Renewable Energy Research Application, Palermo, Italy* (pp. 199–202).
25. Abinayaa, V., & Jayan, A. (2014). Case study on comparison of wireless technologies in industrial applications. *International Journal of Scientific and Research Publications*, 4(2), 1–4.
26. Gagliarducci, M., Lampasi, D. A., & Podesta, L. (2007). GSM-Based monitoring and control of photovoltaic power generation. *Journal of Measurement, Elsevier*, 40(3), 314–321.
27. Smiadak, D. M., & Sözen, M. (2010). Program development for monitoring and analyzing the data from a solar PV system. In *ASEE North Central Sectional Conference, Pittsburgh, PA*.
28. Saranya, K., Reminaa, R.S., & Subhitsha, S. (2016). Modern applications of QR-Code for security. In *2nd IEEE International Conference on Engineering and Technology (ICETECH), Coimbatore, TN, India*.
29. Mohamed, K., Sidi, F., & Jabar, M. A. (2016). Protecting wireless data transmission in mobile application systems using digital watermarking technique. *Journal of Theoretical and Applied Information Technology (JATIT)*, 83(1), 52–63.
30. Liu, Y., & Liu, M. (2016). Automatic recognition algorithm of quick response code based on embedded system. In *IEEE Proceedings of the 6th International Conference on Intelligent Systems Design and Applications (ISDA'06), Jinan, China* (pp. 783–788).

31. Gaikwad, A., & Singh, K. R. (2015). Information hiding using image embedding in QR codes for color images: A review. *International Journal of Computer Science and Information Technologies, (IJCSIT)*, 6(1), 278–283.
32. Thirananant, N., & Lee, H. (2014). A Design of e-Healthcare Authentication Framework with QR Code. *International Journal of Security and Its Applications*, 8(3), 79–86.

# Effects of Activation Function and Input Function of ANN for Solar Power Forecasting



Isha, Akash Singh Chaudhary and D. K. Chaturvedi

**Abstract** Artificial Neural Networks are being used in many applications and forecasting is one of such application where it solves the purpose like stock market predictions, sales forecasting, etc., over the past. In this paper, ANN models are used for forecasting solar power. Multilayer perceptron (MLP) neural network models have been tested for different combinations of transfer functions and net input function on different number of neurons and layers for forecasting solar power. The evaluation and implementation of models are being measured by mean square error.

**Keywords** Artificial neural network · Multilayer perceptron · Solar panel · Transfer function · Solar power forecasting

## 1 Introduction

As we know that the energy consumption of the world is estimated to be greater than twice by the year 2050 and thrice till the century ends. It is much needed that improvements must be done in current energy scenario because the current energy networks are not enough in fulfilling this energy demand required in the coming years in a sustainable way. So it is important to find out substitutes of clean energy for the future. Solar forecasting solves these challenges. It is important to know the factors and reasons which trigger the solar power forecasting the most. These factors include prior idea of the path followed by the sun, the atmospheric conditions, the process of scattering, and the properties of solar plant, i.e., solar panels which contributes the most in creation of solar power by utilizing sun's energy. It is the PV

---

Isha (✉) · A. S. Chaudhary · D. K. Chaturvedi  
Faculty of Engineering, Department of Electrical Engineering, Dayalbagh Educational Institute,  
Agra, India  
e-mail: [ishasingh268@gmail.com](mailto:ishasingh268@gmail.com)

A. S. Chaudhary  
e-mail: [akashsinghchaudhary@gmail.com](mailto:akashsinghchaudhary@gmail.com)

D. K. Chaturvedi  
e-mail: [dkc.foe@gmail.com](mailto:dkc.foe@gmail.com)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_31](https://doi.org/10.1007/978-981-15-0694-9_31)



panels that participate in the conversion of solar energy to electric power. As mentioned earlier the radiations falling on the solar panel and solar panel characteristics decides the output power. There is a sudden rise in the PV power production industry because of the increasing number of users registering everyday. Prior solar forecast knowledge is essential for many reasons like electric grid management, solar power merchandising, etc. As the use of large-scale grid-connected PV system is increasing, it is important to strengthen the prediction of PV system power output, which can help the dispatching department to make overall arrangements for conventional power and photovoltaic power coordination, scheduling adjustment, operation mode planning [1]. There is a significant change in the current Indian power sector which reformulated the perspective and attitude of industries in India. Sustained economic growth continues to drive electricity demand in India. The initiative taken by the Government of India named “Power for all” has acted as a catalyst for solar power industries in the country. At the same time, the competitive intensity is increasing at both the market and supply sides.

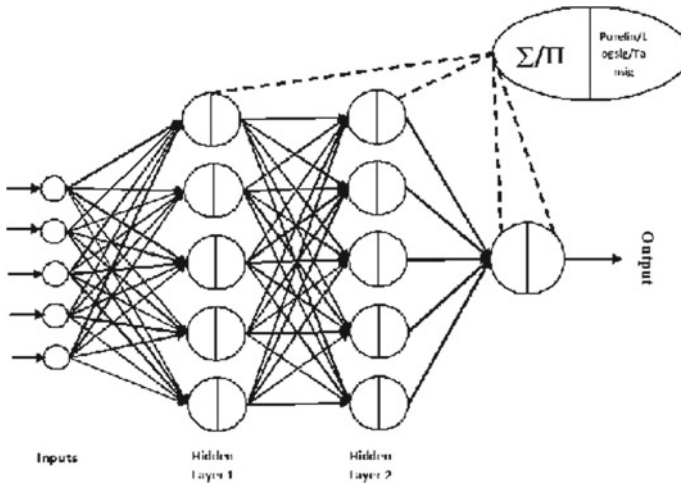
The motivation behind using solar forecasting research is to enhance the grid delivering quality of energy, planning of appliances, and minimization of additional cost associated with weather dependency. The different types of forecasting technique studied and used by researchers are based solely on the fact that where it is applied and executed and for what time period and on what time scale. Among the various, one method of forecasting is through ANNs. Many studies focused on forecasting solar irradiance through ANNs [2, 3]. Al-Alawi and Al-Hinai [4] used climatological variables as inputs to an ANN. Conventional ANN models use only a particular type of activation and aggregation functions in their neural network models. In this paper, we are using different permutation and combination activations and aggregation functions in our model.

The sole purpose to perform this study is to investigate and record the results of Artificial Neural Network for solar power forecasting for different combinations. ANN network having two hidden layers and one output layer with five input parameters like power, temperature, humidity, wind velocity, and pressure for forecasting solar power is shown in Fig. 1. The effect of change in error is measured on the basis of the number of neurons, number of layers, and transfer function on different layers such as hidden and output.

## 2 Background

### 2.1 Neural Network

When understanding a neural network it is important to know what neural networks really are. A neural network can be defined as a large processing network comprising of processing units and having the natural capability where experience-based data can be stored for future purpose. An artificial neural network (ANN) is a type of



**Fig. 1** A multilayer perceptron network

artificial intelligence technique which resembles the functionality of human brain [5]. It has elements like that of a human neural system, which comprises of cell body, axon, and dendrites. Neuron is considered as the integral functional unit of a human brain. Cell body, axon, and dendrites are the three main regions of a neuron. Axons act as a channel (fibers serving as transmission lines) for dendrites to receive information from neurons. The axon dendrite contact organ is called a synapse. The signal reaching a synapse and received by dendrites are electrical impulses. In particular, neural networks are *nonlinear* modeling techniques [6] that learn by example. In this, data is collected by the user, training is invoked so as to allow the structure of data to automatically learn. ANN too comprises of elements known as neurons which communicates via weight connections to interact with each other. Inputs to artificial neural network are multiplied by corresponding weights. All the weighted inputs are then segregated and then subjected to nonlinear filtering to determine the state or active level of the neurons. In ANN, the configuration of neurons is regular and have highly interconnected topology. There lie one or more layers between input and output layers of a network. There is no standard method or formula or hard and fast rule for deciding inputs, network topology, and training method of ANN. Hence, it is not at all easy, consumes much time, and computer intensive to build an ANN. However, these are real time usable due to inherent parallelisms and noise immunity characteristics.

## 2.2 Activation Functions

The ANN performs the task of summing up the product of the input signal and the associated weight for producing an output or activation function.

For the input unit, activation function behaves as an identity function. A mathematical representation of the relationship between input and output in terms of spatial or temporal frequency is known as a transfer function or network function [7, 8]. The transfer functions mostly depict sigmoidal shape and can take the nonlinear form or piecewise linear form or step function form [9]. They are differentiable, continuous, bounded and increases monotonically. A wide range of transfer functions are there but for our research we have taken the three most used transfer function, i.e., log-sigmoid, tan-sigmoid, and linear transfer functions.

The widely used transfer function is the log-sigmoid transfer function (LOGSIG). The beauty of this is that it takes the input value between  $+$  and  $-$  infinity and gives the output between the range of 0 to 1. The log-sigmoid transfer function is mostly used in multilayer networks that are trained using the backpropagation algorithm because this function is differentiable [10]. In cases or examples where exact shape of the transfer function has less weightage as compared to the speed in such cases hyperbolic tangent transfer function (TANSIG) comes into the picture. This transfer function has an output range of  $-1$  to  $+1$  and has relation with bipolar sigmoid and has the equivalency of  $\tanh(n)$  mathematically. The only difference between the two is that it runs faster than  $\tanh$ , and results show minute numerical differences. It is observed that real-world models mostly show nonlinear input and output characteristics. There are also some models that show the characteristics similar to linear characteristics only when they are operated within a certain range and parameters. Purelin Transfer function is an example whose input and output behavior is within a range of acceptance in these kinds of situation.

## 2.3 Aggregation Function

These input functions or aggregation function calculate a layer's net input by combining its weighted inputs and biases. The MATLAB functions *netsum* and *netprod* are used here.

## 3 Model Design and Description

While designing an ANN model a number of systematic steps are followed. These steps include data collection, building the network, training, and testing the model.

### 3.1 Data Collection

In this work, average power (W), temperature (°C), humidity (%), wind velocity (km/hr), and pressure (hPa) from daily data are used during the training and testing of the NN. One month data is collected after acquisition of power output of solar panel prepared in the Dept. of Electrical Engineering, Faculty of Engineering, Dayalbagh Educational Institute (Deemed University), Agra, India. The output power used is recorded at the interval of every 5 min from 6.00 am to 6.00 pm.

### 3.2 Network Building and Training

During this point, various parameters of the network are specified like the number of hidden layers, number of neurons each layer have, transfer function used in different layers, training function, weights/bias learning function, performance metric used, etc.

For the training, the past or previous three values of the data are considered and applied as input and the fourth value of the data is taken as output. The training pattern is consisting of input vector IN and output vector OUT.

$$\text{Input vector IN} = [IN_1(t)IN_1(t - T)IN_1(t - 2T), IN_2, IN_3, IN_4, IN_5]$$

$$\text{Output vector OUT} = [IN(t + T)]$$

$$\text{Training Pattern} = [IN \text{ OUT}]$$

In this work a comparison is been made on the basis of different number of neurons, different number of layers, corresponding transfer function of hidden layers, different layer input functions (aggregation function), and the corresponding mean square error. Levenberg–Marquardt learning algorithm is used for network training.

### 3.3 Testing the Network

For the purpose of testing the performance of the network, a new data is exposed to the model. 18% new sample solar data has been used and applied for testing purpose.

For network performance evaluation training and testing error are observed.

### 3.4 Neural Network Models

Various neural network models were designed for the purpose of comparisons to be made. A two hidden layers neural network is built and comparison is made based

on different combinations of activation functions mentioned below in Table 3. The models are defined on the combinations made on aggregation functions as detailed in Table 1. The number of neurons in hidden layers are kept as 5, 10, 15, and 20. The results of these eight models are shown in Figs. 2, 3, 4, 5, 6, 7, 8, and 9.

$$X_d = \text{Input pair}$$

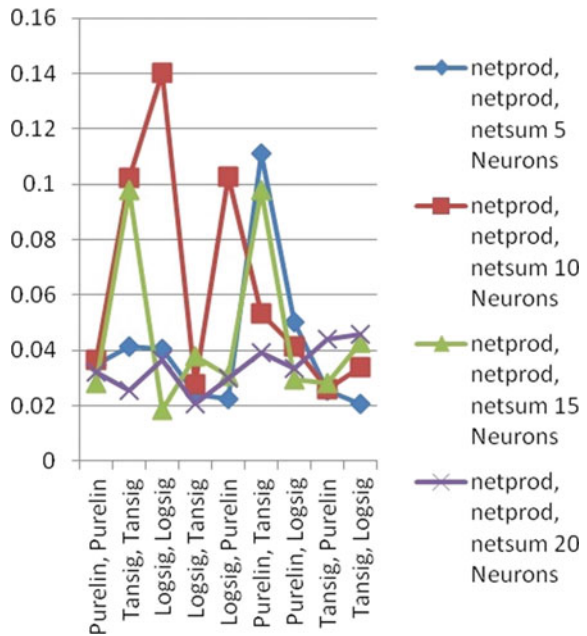
$$d = \text{Number of net}$$

$$j = \text{hidden unit}$$

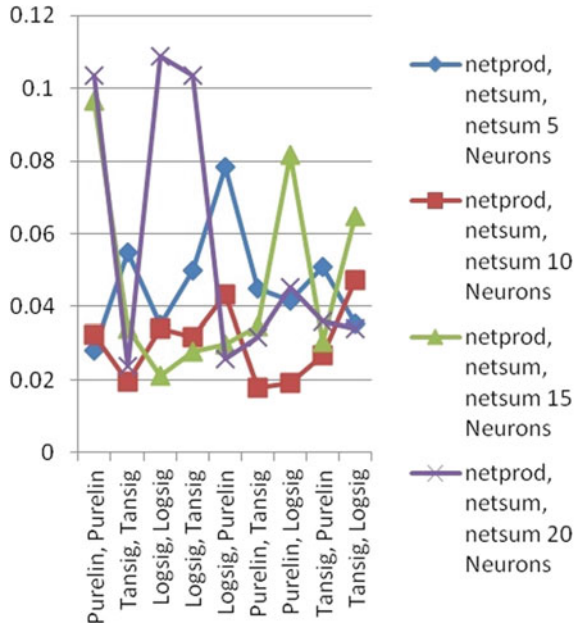
**Table 1** Different neural models for three-layer network

Models	Hidden layer 1 Aggregation function	Hidden layer 2 Aggregation function	Output layer Aggregation function
Model-1	netprod	netprod	netsum
Model-2	netprod	netsum	netsum
Model-3	netsum	netprod	netsum
Model-4	netsum	netsum	netprod
Model-5	netsum	netsum	netsum
Model-6	netprod	netprod	netprod
Model-7	netprod	netsum	netprod
Model-8	netsum	netprod	netprod

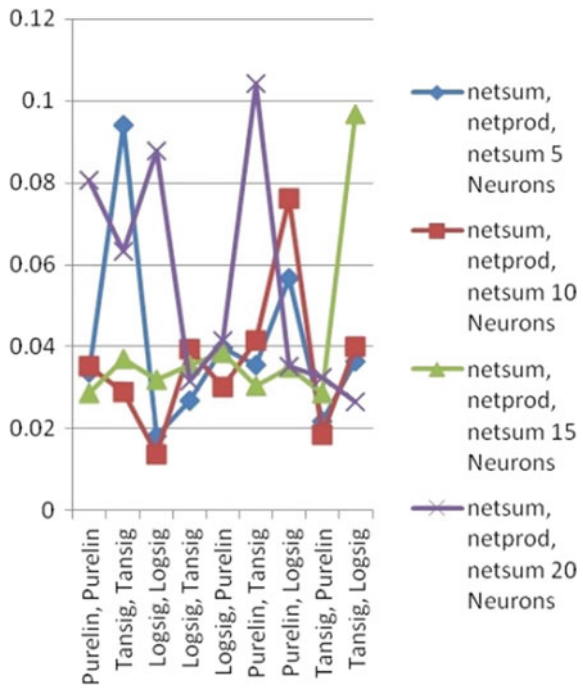
**Fig. 2** Model 1



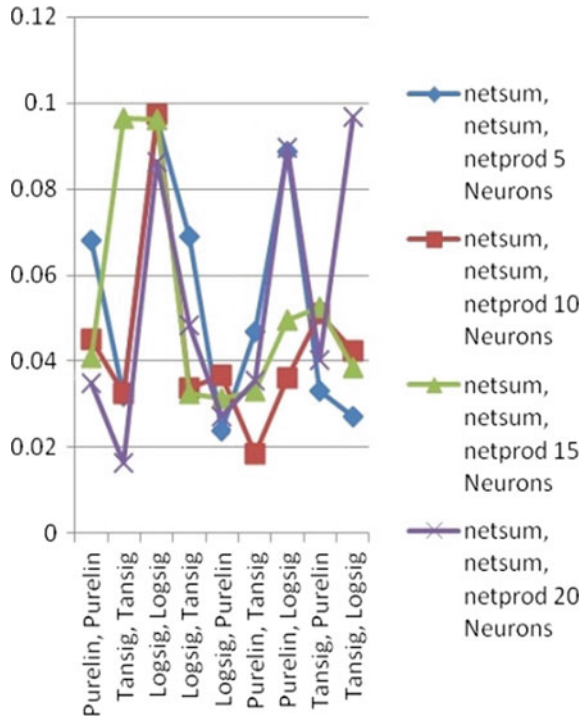
**Fig. 3** Model 2 for three-layer network



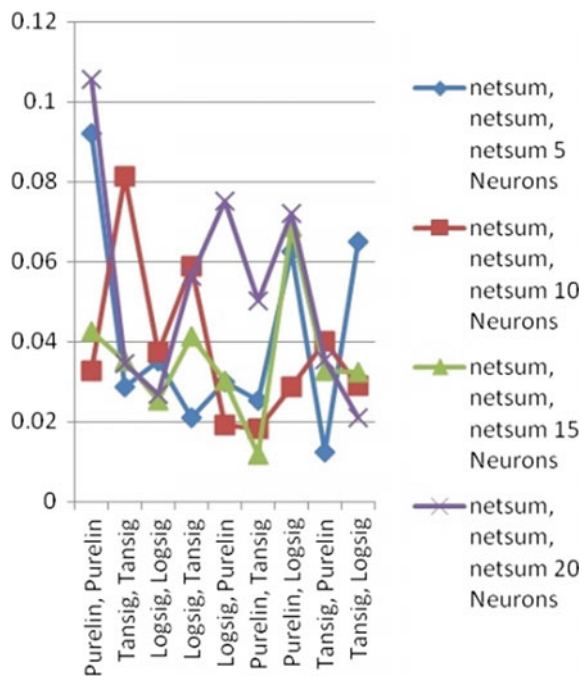
**Fig. 4** Model 3 for three-layer network



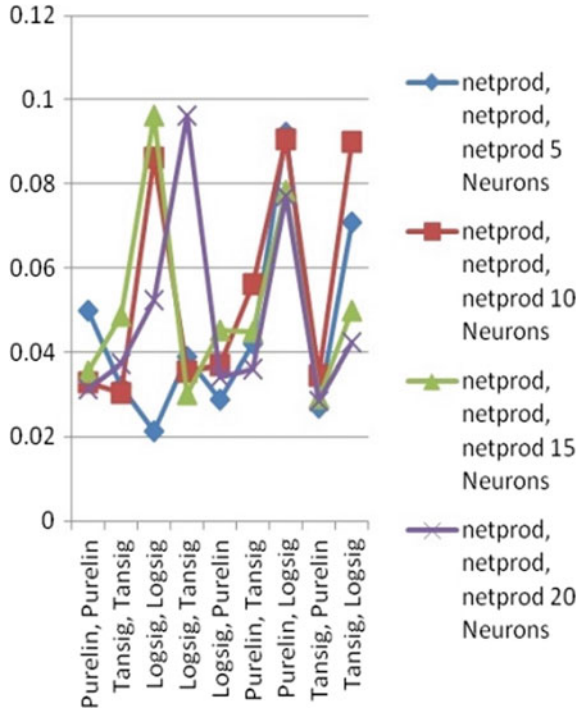
**Fig. 5** Model 4 for three-layer network



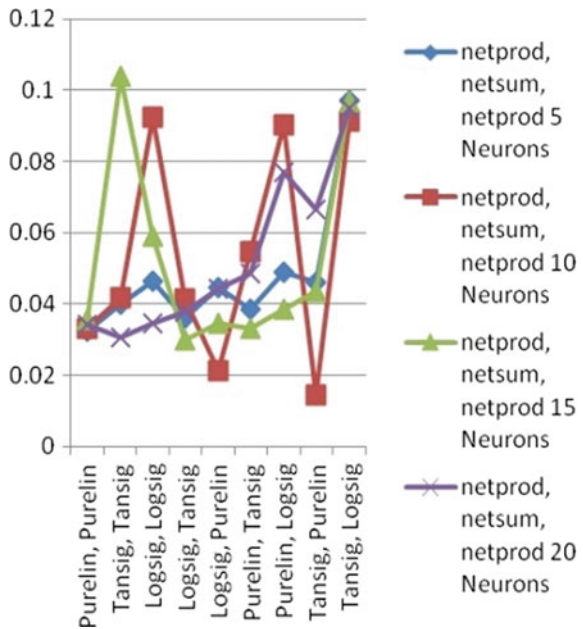
**Fig. 6** Model 5 for three-layer network



**Fig. 7** Model 6 for three-layer network

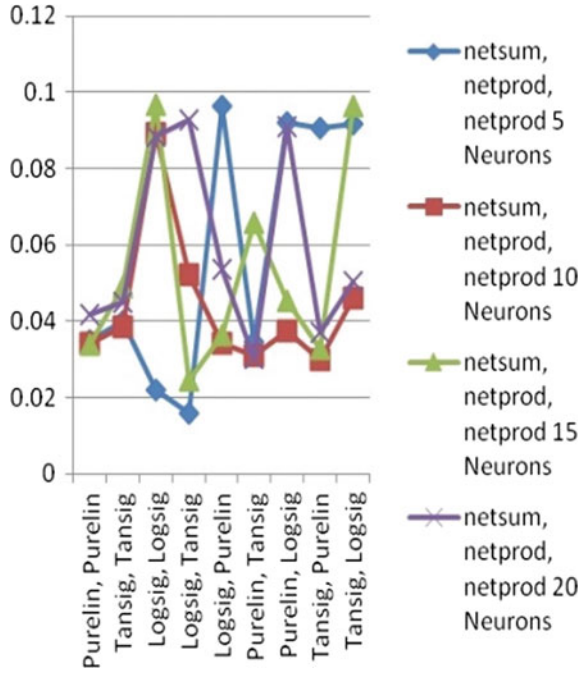


**Fig. 8** Model 7 for three-layer network





**Fig. 9** Model 8 for three-layer network



**Model-1**

$$\text{net}_j^d = \sum_{k=1}^5 w_{jk}x_k^d$$

$$V_{j1}^d = f(\text{net}_j^d) = f\left(\prod_{k=1}^5 w_{jk}x_k^d\right)$$

$$V_{j2}^d = f(\text{net}_{j1}^d) = f\left(\prod_{k=1}^5 w_{jk}x_k^d\right)$$

$$\text{net}_i^d = \sum_{j=1}^5 w_{ij}V_j^d = \sum_{j=1}^5 \left( w_{ij} \cdot f\left(\prod_{k=1}^5 w_{jk}x_k^d\right) \right)$$

$$O_i^d = f(\text{net}_i^d) = f\left(\sum_{j=1}^5 w_{ik}V_j^d\right) = f\left(\sum_{j=1}^5 \left( w_{ij} \cdot f\left(\prod_{k=1}^5 w_{jk}x_k^d\right) \right) \right)$$

Mathematically, model 1 is defined above. Similarly, other models can also be defined (Table 2).

We have also recorded the results for one hidden layer network, i.e., a neural network with inputs, one hidden layer, and one output layer. Four models are defined

**Table 2** Activation function combinations for three-layer network

Three-layer network	Hidden layer 1 Activation function	Hidden layer 2 Activation function	Output layer Activation function
Activation function combinations	Purelin	Purelin	Purelin
	Tansig	Tansig	Purelin
	Logsig	Logsig	Purelin
	Logsig	Tansig	Purelin
	Logsig	Purelin	Purelin
	Purelin	Tansig	Purelin
	Purelin	Logsig	Purelin
	Tansig	Purelin	Purelin
Tansig	Logsig	Purelin	

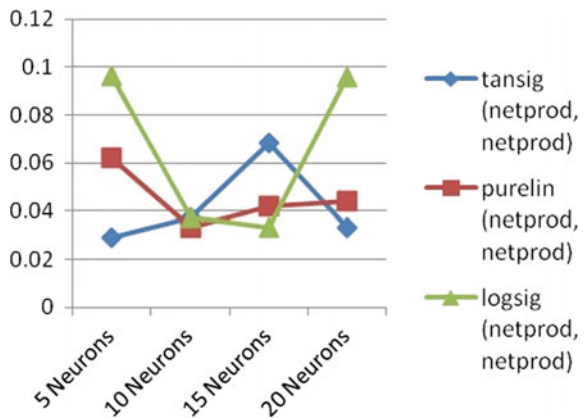
**Table 3** Different Neural Models for two-layer network

Models	Hidden layer 1 Aggregation function	Output layer Aggregation function
Model-1	netsum	netprod
Model-2	netsum	netsum
Model-3	netprod	netprod
Model-4	netprod	netsum

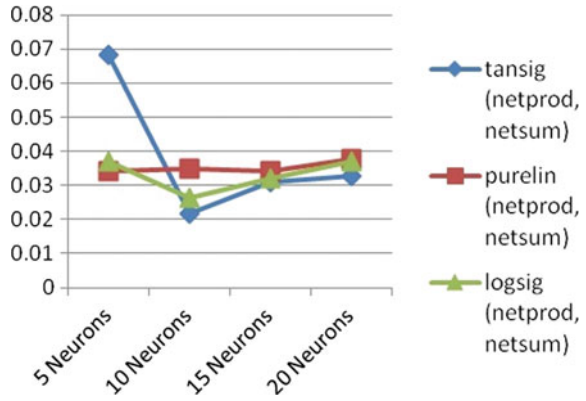
in this case as given in Table 3. The results of these models are shown in Figs. 10, 11, 12, and 13.

The combinations of transfer functions for two-layer network are mentioned in Table 4.

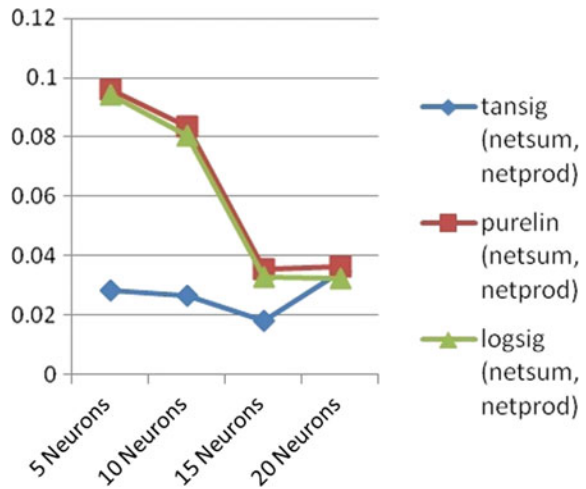
**Fig. 10** Model 1 for two-layer network



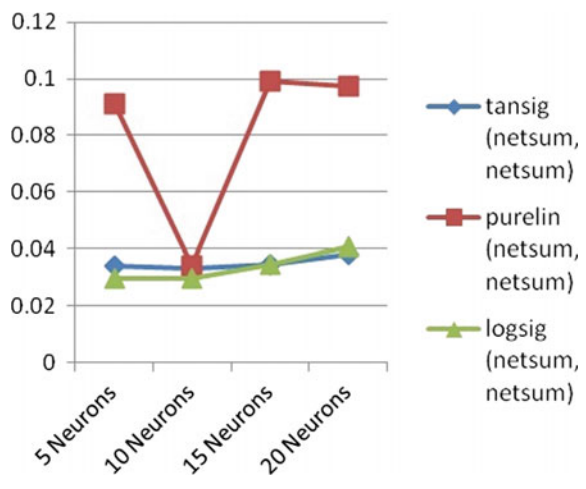
**Fig. 11** Model 2 for two-layer network



**Fig. 12** Model 3 for two-layer network



**Fig. 13** Model 4 for two-layer network



**Table 4** Activation function combinations for two-layer network

Two-layer network	Hidden layer 1 Activation function	Output layer Activation function
Activation function combinations	Tansig	Purelin
	Purelin	Purelin
	Logsig	Purelin

### 3.5 Results and Discussion

This section presents the best results achieved through ANN models. The graphs below show the computed mean square error for different combinations activation function for different number of neurons and different layer input functions. The results are discussed for 50 epochs and performance goal is kept as 0.0000001.

The results are discussed for three-layer network where different combinations of activation functions are applied on two hidden layers. The activation function for output layer is kept purelin for all the combinations. The combination of input functions is also observed for three layers, i.e., two hidden and one output. Three-layer network MSE graphs are shown by Figs. 2, 3, 4, 5, 6, 7, 8, and 9. Results are also observed for two-layer network having one hidden and one output layer. In this case also the activation function for output layer is kept as purelin, whereas hidden layer activation functions are changed. Two-layer network MSE graphs are shown in Figs. 10, 11, 12, and 13.

The best results are given by Purelin, Tansig activation function for 15 numbers of neurons when all the three layers, i.e., two hidden and one output have input function (or aggregation function) as *netsum*. When different combinations are tested for different layers it is found that *netsum*, *netprod*, *netsum* combination for two hidden layers and one output layer, respectively, gives best result for 10 numbers of neurons when activation functions are Logsig, Logsig for the two hidden layers. It can be observed from the simulations that as the number of neurons are increased the error decreases. The highest error is shown by five neurons for Logsig, Logsig activation function when all the three layers have input function as *netprod*. When comparison is made between three-layer network and two-layer network it is found that error decreases on increasing the number of layers. For two-layer network the minimum error is recorded for 15 numbers of neurons having hidden layer and output layer activation function as tansig and purelin, respectively. The input function combination for hidden layer and output layer is kept as *netsum*, *netprod*, respectively.

## 4 Conclusions

The paper shows the comparison between three-layer network and two-layer network. Different transfer function combinations for two hidden layers in three-layer

network and one hidden layer in two-layer network are compared for different layers combination of input function for forecasting solar power. It is observed that feed forward neural network with Purelin, Tansig activation function for 15 number of neurons when all the three layers have input function as *netsum* shows the best result. This model has the minimum error. The Levenberg–Marquardt training algorithm is used for ANN forecasting model.

## References

1. Ding M., Wand L., & Bi. R. (2011). An ANN-based Approach for Forecasting the power output of photovoltaic system. *Procedia Environmental Sciences*, 11,1308–1315.
2. Sozen, A., Arcaklioglu, E., Ozalp, M., & Caglar, N. (2005). Forecasting based on neural network approach of solar potential in Turkey. *Renewable Energy*, 30(7), 1075–1090.
3. Mellit, A., & Pavan, A. M. (2010). A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a gridconnected PV plant at Trieste, Italy. *Solar Energy*, 84(5), 807–821.
4. Al-Alawi, S., & Al-Hinai, A. (1998). An ANN- based approach for predicting global solar radiation in locations with no measurements. *Renewable Energy*, 14(1–4), 199–204.
5. Haykin, S. (2009). *Neural networks and learning machines* (3rd ed.). New Jersey: Pearson Education Inc.
6. Chaturvedi, D. K. (2010). *Modeling and simulation of systems using MATLAB and simulink*. CRC Press.
7. Agatonovic-Kustrin, S., & Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of Pharmaceutical and Biomedical Analysis*, 22(5), 717–727.
8. Yitian, L., & Gu, R. R. (2003). Modeling flow and sediment transport in a river system using an artificial neural network. *Environmental Management*, 31(1), 122–134.
9. Wilby, R. L., Wigley, T. M. L., Conway, D., Jones, P. D., Hewitson, B. C., Main, J., et al. (1998). Statistical downscaling of general circulation model output: A comparison of methods. *Water Resources Research*, 34(11), 2995–3008.
10. Dorofki, M., Elshafie, A. H., Jaafar O., Karim O. A., & Mastura, S. (2012). Comparison of Artificial Neural Network Transfer Functions Abilities to Simulate Extreme Runoff Data. In *International Conference on Environment, Energy and Biotechnology*, 33, 39–44.

# An Integrated Approach Toward Smart Parking Implementation for Smart Cities in India



Ishan Kumar, Prashant Manuja, Yashpal Soni and Narendra Singh Yadav

**Abstract** This paper discusses an integrated approach toward smart parking implementation for smart cities in India using radar detection and ultrasonic technology. The aim of this research is to develop an autonomous parking system which could reduce traffic congestion and toxic emission from vehicles and also make it easier for customers to find a place to park during peak hours, hence saving their time. The autonomous parking system will have minimum amount of human interaction, and all the data will be sent directly to a cloud server wirelessly using a Wi-Fi module. In addition to that, the interface will also guide the customer to an available parking space. The paper discusses a system that can perform real-time monitoring of parking space availability by individual space and can be embedded in software. The software will raise a system alert when the number of vehicles in transit and more exceeds. The project can be implemented in various situations in places like schools, colleges, institutions, etc., and can provide an end-to-end solution for a safe parking mechanism with low-cost maintenance.

**Keywords** Integrated approach · Smart parking · Radar detection · Autonomous parking system · End-to-end solution

---

I. Kumar (✉) · P. Manuja · Y. Soni · N. S. Yadav  
Manipal University Jaipur, Jaipur 303007, India  
e-mail: [ishanlko2001@gmail.com](mailto:ishanlko2001@gmail.com)

P. Manuja  
e-mail: [prashant.manuja@jaipur.manipal.edu](mailto:prashant.manuja@jaipur.manipal.edu)

Y. Soni  
e-mail: [yashpal.soni@jaipur.manipal.edu](mailto:yashpal.soni@jaipur.manipal.edu)

N. S. Yadav  
e-mail: [narendrasingh.yadav@jaipur.manipal.edu](mailto:narendrasingh.yadav@jaipur.manipal.edu)

## 1 Introduction

A **smart city** is a municipality that uses information and communication technologies to increase operational efficiency, share information with the public, and improve both the quality of government services and citizen welfare [1]. A study was conducted by INRIX which states that 20% of traffic in urban areas can typically be attributed to people searching for parking. Most of the existing outdoor parking sensors that are currently on the market have accuracy limitations, which can negatively impact a person's parking experience. Previous technologies also struggled with wireless capabilities and interference from cellular networks commonly associated with urban environments.

The traditional parking system revolves around the supply and demand in a city environment. With the growing pace of vehicle usage that is happening across cities, the parking problem is growing fourfold, and this cannot be managed by way of physical security guards who monitor this day in and day out. Cost and customer ease are compromised on a daily basis, and the present system leads to dissatisfaction in the citizens of these cities.

When parking, it is a quest to find a free spot, especially during peak hours, causing any individual to drive around for hours on end to just find one. Such situations can waste a lot of time resulting in the unnecessary amount of toxic emission, extensive traffic, and impatient drivers.

## 2 Integrated Approach for Smart Parking

Each parking space is equipped with a centrally powered occupancy sensor that can detect the absence, arrival, presence, and departure of a vehicle. The sensors are connected to NodeMCU Wi-Fi Module having its own unique 32-bit service set identifier (SSID) and password which will be connected to dedicated server which holds all the details of the occupancy. While the occupancy sensor consists of two components: proximity sensor and the ultrasonic radar detection system. Both of these contribute to the detection of the vehicle: the first one detects the presence of an object while the other confirms it as a vehicle so this way, we can be double sure about the presence of the vehicle.

When an occupancy sensor detects vehicle activity, both the sensors send a digital signal to the NodeMCU module which is then converted into packets and sent over a wireless network with a unique SSID, which ensures that the data being sent to the correct location. Now the server collects the data and uses the sensors' messages to keep track of open and occupied spaces and display then on a Central Screen, which makes it more user-friendly. The working of a single occupancy sensor is shown in Fig. 1.

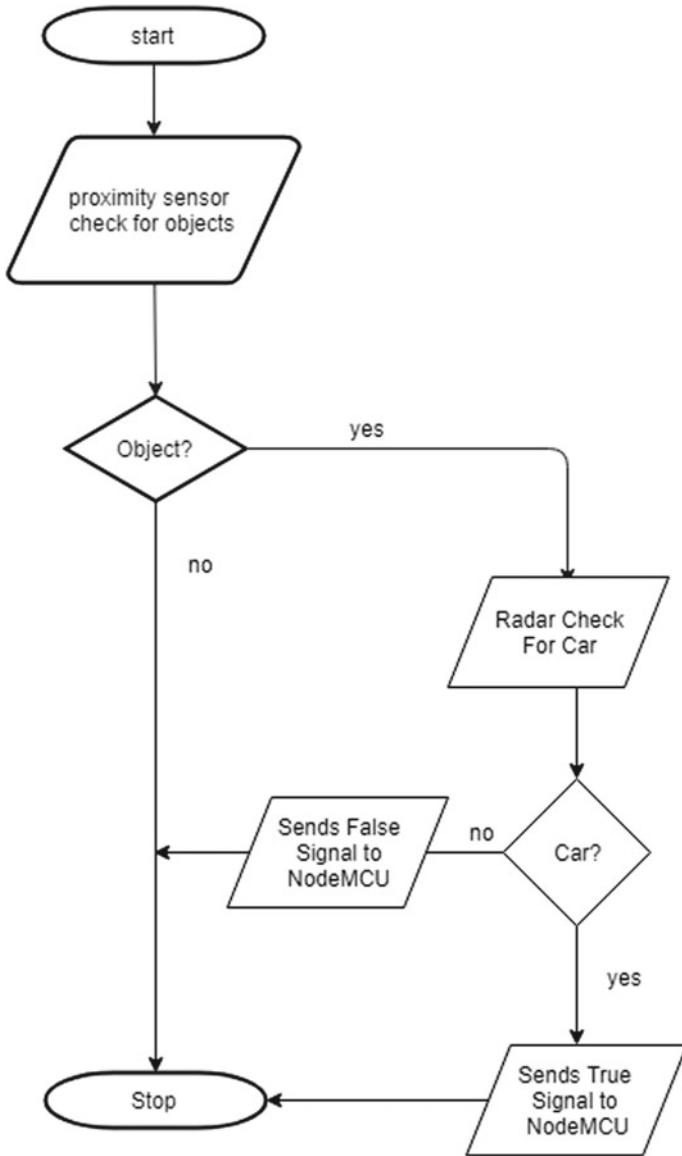


Fig. 1 Flowchart of proposed solution



## 3 Hardware Description

### 3.1 *NodeMCU*

**NodeMCU** is an open-source IoT Platform [2]. It is the motherboard of the whole sensor system. It consists of firmware which runs on the ESP2866 Wi-Fi module and hardware which is based on ESP-12 module. Both the sensors send the digital signal to this board which is then sent to the server wirelessly [3].

### 3.2 *Occupancy Sensor*

The occupancy sensor consists of two components.

#### 3.2.1 **Ultrasonic Radar System**

This consists of an HC-SR04 ultrasonic distance measurement sensor mounted on a servo motor which acts as an axis of rotation for the system. When a car is parked in a vacant occupancy, the sensor calculates the distance between the vehicle and the parking boundary; if the distance lies in a certain range of value, this means that the car is parked properly in the occupancy, and the system sends a digital signal to the nodeMCU module.

#### 3.2.2 **Optical Proximity Sensor**

The ultrasonic sensor used in the radar system acts as a proximity sensor and takes reading of the obstacles continuously in a timespan of 2 ms and sends digital signal as an output. When a car is parked in occupancy, it detects the presence of a vehicle, and then, it gives a digital signal to the nodeMCU module.

## 4 Implementation

The sensor is placed in front of the parking space and then following things is done.

### 4.1 *Approximation of Theta*

Car detection through radar is done simply by using the arc length formulae.

**Table 1** Width of cars

Car name	Width (in meters)
Maruti Suzuki Wagon R	1.475
Hyundai Santro	1.645
Maruti Suzuki Dzire	1.735
Mahindra XUV500	1.89
Toyota Fortuner	1.85
Jeep Compass	1.818

Source <https://auto.ndtv.com/>

**Table 2** Theta vs Arc length

Theta ( $\theta^\circ$ )	Arc length (in meters)
60	1.04
70	1.22
80	1.39
90	1.57
100	1.74
110	1.91
120	2.09

That is:

$$L(\text{arc length}) = \frac{\theta^\circ}{360^\circ} 2\pi r$$

The width of various cars is given above (Table 1).

For this system to work properly with all the cars, the minimum width should be taken that is 1.47 m.

Therefore, the arc length covered by the radar should be greater than 1.47 m. So testing was done by changing the value of  $\theta$  from  $60^\circ$  to  $120^\circ$  with  $r$  constant at 1 m and calculated the arc length.

From Table 2, it is clear that the optimum value of theta in order to detect any type of car is  $90^\circ$ .

## 4.2 Detection of Car-like Object

Since the perpendicular distance between the car and radar is 1–1.5 m, we calculated the distance at the extreme points of the car that is at  $45^\circ$  both sides of the radar when its perpendicular distance between the car and radar is 1–1.5 m.

From Table 3, it is very clear that for this system to work properly for all the cars, the maximum distance has to be considered, that is 2.12 m.

**Table 3** Perpendicular distance vs Extreme point distance

Perpendicular distance (in m)	Extreme point distance (in m)
1	1.41
1.1	1.55
1.2	1.69
1.3	1.83
1.4	1.97
1.5	2.12

Now the sensor calculates the distance after every 5° of rotation and checks whether the calculated distance is less than 2.12 m or not, if not it increases a flag value by 1. After one complete time period, the value of flag is checked. If it is less than 2, then the radar sends a true signal to the nodeMCU, which means that the object is a car; otherwise, it sends a false signal.

## 5 Comparison Between Various Technologies

Previous papers include the use of geographic location sensor [4], which increases the complexity and decreases the ease of use, while the project discussed in this paper is easy to use since all the sensors are connected to a single server and every occupancy sensor can itself respond to the server. This technique decreases the latency in data transfer using NodeMCU.

Cost efficiency is another aspect in which this project has an edge over the others.

Since all the work is done wirelessly, fewer components are used so production and installation cost cuts to about half when compared to companies who use RFID chips and geomagnetic sensors [5].

## 6 Conclusion

The usage of the above approach will help the market place utilize the technology as it will develop an autonomous parking system. This will reduce traffic congestion and toxic emission from vehicles and also make it easier for customers to find a place to park during peak hours, hence saving their time, as a result of its ease of usage in the parking solutions segment. The implementation itself will take less time in comparison to other models which exist as of today. Through this approach, customer can perform real-time monitoring of parking space availability by individual space from anywhere in the world.

## References

1. <https://internetofthingsagenda.techtarget.com/definition/smart-city>.
2. <https://en.wikipedia.org/wiki?curid=46194161>.
3. Anjari, L., & Budi, A. H. S. (2018). The development of smart parking system based on NodeMCU 1.0 using the internet of things. In *IOP Conference Series: Materials Science and Engineering* (Vol. 384, p. 012033). <https://doi.org/10.1088/1757-899x/384/1/012033>.
4. Petsch, K., Dotzlaw, P., Daubenspeck, C., Duthie, N., & Mock A. (2012) Auto mated parking space locator: RSM. In *Proceedings of the 2012 ASEE North Central Section*.
5. Pala, Z., & Inanc, N. (2007). Smart parking applications using RFID technology. In *2007 1st Annual RFID Eurasia* (pp. 1–3). <https://doi.org/10.1109/rfideurasia.2007.4368108>.

# Distributed Processes Scheduling Based on Evolutionary Approach



Santosh Kumar, Gaurav Dubey and Shailesh Tiwari

**Abstract** The main aim of the distributed system is to maximize the utilization of resources with minimal response time and overall execution time. For this, optimal scheduling of tasks is desirable, which is also a well-known NP-complete problem. This paper presents an algorithm to generate an optimal schedule for distributed system considering the parameters like processor load, peak load, average processor utilization, and communication cost. Further, algorithms have been experimentally compared, and the proposed algorithm has performed better than existing algorithm.

**Keywords** Genetic algorithm · Distributed system · Processor utilization · Communication delay

## 1 Introduction

The distributed system is a collection of interconnected autonomous heterogeneous computing machines and coordinated by a software entity called distributed operating system. The main aim of the distributed system is to maximize the utilization of resources with minimal response time and overall execution time. For this, optimal scheduling of tasks is desirable which is also a well-known NP-complete problem [1]. In distributed task scheduling problem, priority of tasks plays an important role for both heuristic and search algorithms, and it influences the scheduling results overall execution time. The scheduling in distributed system is a global issue which is done in coordination with other nodes. Process can be migrated to other nodes to balance the system load. By transferring the processes from more busy processor to another

---

S. Kumar (✉) · G. Dubey · S. Tiwari  
Department of Computer Science and Engineering, ABES Engineering College, Ghaziabad  
201009, India

e-mail: [santoshg25@gmail.com](mailto:santoshg25@gmail.com)

G. Dubey  
e-mail: [gdubey1977@gmail.com](mailto:gdubey1977@gmail.com)

S. Tiwari  
e-mail: [shailesh.tiwari@abes.ac.in](mailto:shailesh.tiwari@abes.ac.in)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_33](https://doi.org/10.1007/978-981-15-0694-9_33)

less busy processor, overall load can be balanced. In distributed environment, each node creates its own strategy to allow or reject tasks.

Scheduling of tasks can be performed in four parts: the selection policy plays a role that which job to be migrated, the transfer policy decides when to migrate, the location policy decided partner node, and the information policy maintains the system state from all nodes [2, 3]. The scheduling can be done either statically or dynamically and is based on the process characteristics and the current system state [4]. Static scheduling algorithms minimize the overall execution time of parallel tasks by minimizing the communication delay [5]. Static scheduling algorithms need to predict execution time and communication delay at compile time. To reduce the communication delay, it prepares coarser-grain processes from smaller tasks. On the other hand dynamic scheduling algorithms redistribute the tasks among the processors during run time from deeply loaded processor to evenly loaded processors. Any load balancing algorithm in dynamic scheduling adopts three policies: information policy defines the amount of load information for job assignment, transfer policy finds the condition such as existing load of the host and the size of the job, and placement policy determines the processor to which a job may be transferred. Thus, the scheduling of task in distributed systems can be defined as how to execute a set of tasks  $T$  on a set of processors  $P$  subject to the constraint of optimization criteria  $C$ . Several algorithms use different location policies such as random location policy, threshold location policy, and shortest location policy. Random location policy: In this location policy, the system uses no remote state information. In this way, select a node randomly and transfer the task to that node, without exchanging any information between the nodes. A task transfer becomes useless when the receiver node highly loaded [6].

## 2 Related Work

In distributed system, load balancing by distributing the tasks among two or more sites, network nodes, CPUs, disks or other resources, is needed for maximizing the resource utilization, throughput, and response time [7, 8]. Balancing of load and task setup is multifarious problem in multiprocessor systems, and this is NP-hard problem [9]. Several distributed process scheduling algorithms have been proposed. These methods can be kept in broadly three categories: graph theory-based approaches [10], mathematical model-based approaches [11], and heuristic-based approaches [12–14]. Since the scheduling problems are NP-complete [15], heuristic techniques can provide an optimal solution. GA-based [14, 16–21] and simulated annealing [21, 22] based approaches proposed in the literature for distributed process scheduling. The main aim of distributed process scheduling is load balancing. Processes are

created randomly in the system so some processors may be heavily overloaded while the others lightly loaded even some processors may be idle. The aim of scheduling processes is to distribute the load on different processors uniformly so that processors utilization could be maximized and total execution time could be minimized [23]. In [24], sender-initiated and receiver-initiated approaches using GA for load balancing in distributed systems have been proposed. Both the approaches result in redundant request messages leading poor system performance. CPU queue length has been used as load index [24–26]. Their approach used three time parameters: total message processing time, total time to transfer messages from the sender to intended processors covering shortest distance, and total task processing time at each processor. In [7], GA is used to address the job scheduling problem by using balancing of load efficiently. Hereafter this algorithm will be referred as NK (Nikravan M. and Kashani M. H.) algorithm. Scheduling of processes is done based on processor load. Performance and efficiency are measured for these schedules.

### 3 Proposed Approach

#### 3.1 Tasks Schedule Representation

A candidate schedule is a set of tasks which also represent the candidate solution. Here  $T_1, T_2, T_3, \dots, T_n$  represent the tasks, and indices represent the process identifier allocated to definite processor. This schedule is generated randomly initially and is shown in Fig. 1:

Total processes,  $n$ .

Total processors,  $m$ .

Find set of top- $m$  tasks (schedule of  $T$  out of  $n$  tasks) to run on  $m$  processors.

#### 3.2 Crossover

Crossover creates a new schedule by exchanging some tasks from two randomly chosen schedules. Here cyclic crossover, shown in Fig. 2, is used since any individual task can not repeat in a schedule.

Tasks	$T_{10}$	$T_2$	$T_{23}$	$T_4$	$T_5$	$T_{16}$	$T_7$	$T_{28}$	$T_{19}$
PID/Chromosome/Schedule(SD)	10	2	23	14	5	16	17	28	19
Processors	P1	P2	P3	.	.	.	.	.	$P_m$

Fig. 1 Tentative process schedule

Before cyclic cross over

<b>CS1</b>	4	6	8	11	1	10	5	3	7	15	9	2
<b>CS2</b>	6	5	3	12	4	8	9	7	10	2	1	11

After cyclic cross over

<b>OS1</b>	4	6	8	11	1	10	5	3	12	9	7	2
<b>OS2</b>	6	5	3	12	4	8	11	1	10	7	15	9

Fig. 2 Cyclic\_Crossover

### 3.3 Mutation

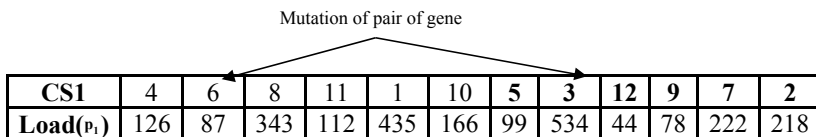
Mutation is applied to provide diversity in successive generations. In some schedules, a few tasks are flipped by some other randomly selected task. In the proposed approach, two processors  $p_u$  and  $p_v$  with minimum and maximum execution time, respectively, are selected. Then their tasks are swapped so that performance is improved, and maximum load is reduced [27], e.g., as shown in Fig. 3.

### 3.4 Fitness Function

A fitness function is used to evaluate a solution which considers parameters like processor load, peak load, communication cost and average processor utilization [7].

#### Input

- Total processes,  $p$ .
- Total processors,  $q$ .
- Initial population size, IPS.
- Total iterations,  $N$ .
- Mutation probability, MP.
- $N1$ —Total processes allocated to processor  $i$ .
- $N2$ —Total number of new processes allocated to processor  $i$ .



Chromosome after mutation

<b>CS1</b>	4	3	8	11	1	10	5	6	12	9	7	2
<b>Load(<math>p_i</math>)</b>	126	87	343	112	435	166	99	534	44	78	222	218

Fig. 3 Mutation



- N3—Total number of newly arrived processes.
- N4—Total number of processors.

Communication delay between processors,  $0 \leq T1[N4][N4] \leq Target\_Value.$

Data transfer time of unit data,  $0 \leq DTT[N4][N4] \leq Target\_Value.$

Execution time,  $0 \leq ET [ p ] [ qm ] \leq Target\_Value.$

Data volume,  $0 \leq DV [ 1 ][ p ] \leq Target\_Value.$

Selected processor,  $0 \leq SP [ 1 ][ p ] \leq Target\_Value.$

Existing processor,  $0 \leq CP [ 1 ][ p ] \leq Target\_Value.$

### 3.4.1 Processor Load

It depends on currently allocated processes and that may be allocated later on [7] and is given below:

$$Load(P_i) = \sum_{j=1}^{NP} ET_{j,i} + \sum_{k=1}^{NNP} ET_{k,i}$$

where

NP = No. of processes allocated to processor i.

NNP = No. of newly allocated processes to processor i.

### 3.4.2 Peak Load

It is maximum load allocated to any processor.

$$PeakLoad(M) = \max(Load(P_i))$$

where M = maximum execution time; Pi = processor with maximum load.

### 3.4.3 Communication Cost

It is delay rate between two processors and transfer of d-size data between them and is given below

$$Communication\ Cost = \sum_{i=1}^{np} (CD_{c_i f_i} + DT_{c_i f_i}^* d_{1,i})$$

### 3.4.4 Average CPU Utilization

Average CPU utilization is based on the utilization of particular processor with respect to total number of processors and is given as below:

$$Utilization(Processor_i) = \frac{Load(Processor_i)}{Peak\_Load} \quad (4)$$

$$Avg\_CPU\_Utilization(Processor_i) = \frac{\sum_{i=1}^{no\_of\_processors} Utilization(Processor_i)}{no\_of\_processors} \quad (5)$$

$$Cost\_of\_Schedule(S) = \frac{Average\_CPU\_Utilization * \lambda}{Peak\_Load * Communication\_cost} \quad (6)$$

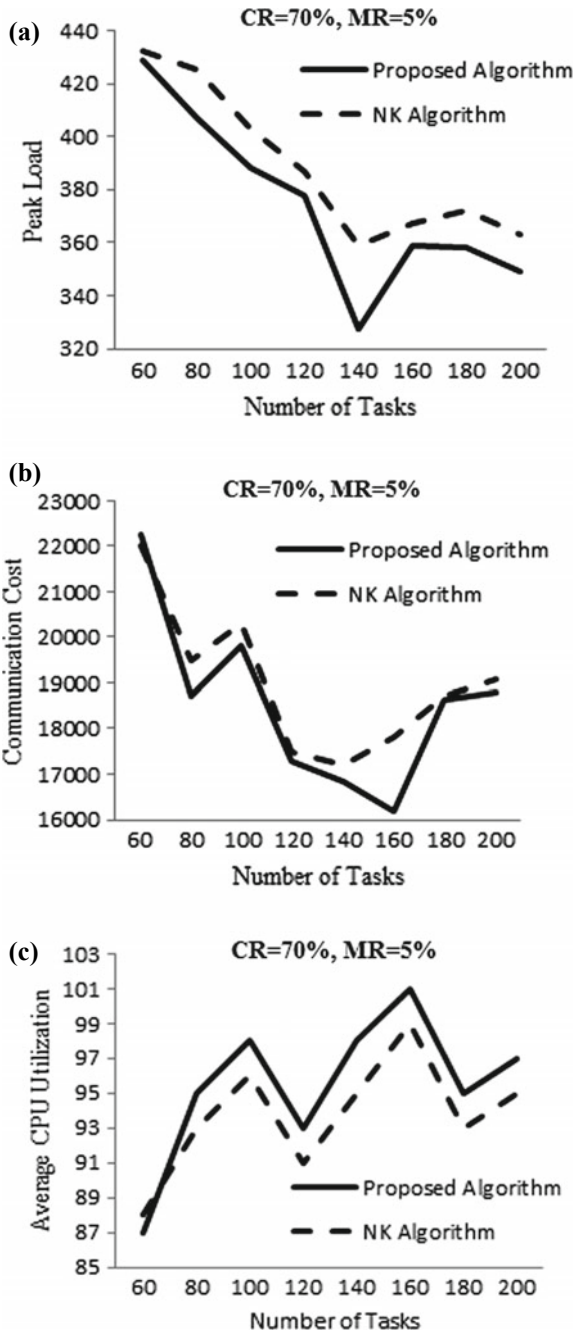
where  $\lambda$  is constant.

Now the proposed algorithm using fitness function of Eq. 6 is given next in Fig. 4. The algorithm selects top-T schedule as Schedule[0].

**Algorithm:**

1. Produce Initial Schedules of randomly generated candidate schedule, Schedule[m]
2. Evaluate each schedule
  - For i=1 to q
  - Cost<sub>i</sub> = fitness(Schedule[i])
  - End for
3. For Generation=1 to G //G is Maximum Generations
  - i. Perform tournament selection
    - i=1
    - while(i<= Crossover\_Rate)
      - random variable, 0 < r < 1
      - CR = Crossover\_Rate / 100
      - if r < CR then
        - Append fitter Schedule to CrossoverSchedule[]
      - otherwise
        - Append less fitter Schedule to CrossoverSchedule[]
  - ii. Randomly choose two schedule P1 & P2 from crossover schedule
  - iii. Perform crossover between Schedule P1 & P2, Generating Offsprings Off1 & Off2
  - iv. Append Off1, Off2 to NewSchedule[]
  - v. Perform mutation with rate MR
  - vi. Schedule[] ← NewSchedule[]
  - vii. Gen=Gen+1
4. return Schedule[0]

**Fig. 4** Proposed algorithm



**Fig. 5** a Peak load versus number of tasks. b Communication cost versus number of tasks. c Average CPU utilization versus number of tasks

## 4 Experimentation and Result

Both the algorithms were implemented in JDK1.7. These algorithms were compared after conducting results. The proposed algorithm performs better on the considered parameters as shown in Fig. 5a, b and c.

## 5 Conclusion

In distributed system, process scheduling plays a key function in by and large system performance and throughput. GA-based approach for generating an optimal schedule has been proposed. Proposed methodology considers several parameters to address distributed scheduling problem and optimizes the peak load and communication cost; also it improves the average CPU utilization and balances the load among different sites. Further, results illustrate that the proposed approach performed superior to the existing approach.

## References

1. Yao, W., Yao, J., & Li, B. (2004). Main sequences genetic scheduling for multiprocessor systems using task duplication. *International Journal of Microprocessors and Microsystems*, 28, 85–94.
2. Chaptin, S. J. (2003). *Distributed and multiprocessor scheduling*. University of Minnesota.
3. Tel, G. (1998). *Introduction to distributed process scheduling*. University of Cambridge.
4. Sarkar, V., & Hennessy, J. (1986). Compile-time partitioning and scheduling of parallel programs. In *Symposium Compiler Construction* (pp. 17–26). New York: ACM Press.
5. Eager, D. L., Lazowska, E. D., & Zahorjan, J. (1986). Adaptive load sharing in homogeneous distributed systems. *IEEE Transactions on Software Engineering*, SE-12(5), 662–675.
6. Chapin, S. J., Weismann, J. B. (2002). Distributed and multiprocessor scheduling. *Electrical Engineering and Computer Science Head book* (pp. 40).
7. Nikravan, M., & Kashani, M. H. (2007). A genetic algorithm for process scheduling in distributed operating systems considering load balancing. In *21st European Conference on Modelling and simulation*, ECMS.
8. Singh, R., & Gupta, S. K. (2013) Distributed process scheduling using genetic algorithm, in the next generation information technology summit. *IET Digital Library and IEEE Xplore*, 48–54. (26th–27th Sept. 2013).
9. Lee S., Lee D., Lee W., & Cho H. (2004). An adaptive load balancing approach in distributed computing using genetic theory. In K. M. Liew, H. Shen, S. See, W. Cai, P. Fan, & S. Horiguchi (Eds.) *Parallel and distributed computing: Applications and technologies* (Vol. 3320). Lecture Notes in Computer Science. Berlin: Springer.
10. Ma, P. Y. R., Lee, E. Y. S., & Tsuchiya, J. (1982). A task allocation model for distributed computing systems. *IEEE Transactions on Computers*, 31(1), 41–47.
11. Park, G. L. (2004). Performance evaluation of a list scheduling algorithm in distributed memory multiprocessor systems. *International Journal of Future Generation Computer Systems*, 20, 249–256.
12. Woodside, C. M., & Monforton, G. G. (1993). Fast allocation of processes in distributed and parallel systems. *IEEE Transactions on Parallel and Distributed Systems*, 4(2), 164–174.

13. Sarje, A. K., & Sagar, G. (1991). Heuristic model for task allocation in distributed computer systems. *Proceedings of the IEE-E*, 138(5), 313–318.
14. Park, C. I., & Choe, T. Y. (2002). An optimal scheduling algorithm based on task duplication. *IEEE Transactions on Computers*, 51(4), 444–448.
15. Shen, C. C., & Tsai, W. H. (1985). A graph matching approach to optimal task assignment in distributed computing using a minimax criterion. *IEEE Transactions on Computers*, 34(3), 197–203.
16. Martino, V.D. (2003). Sub optimal scheduling in a grid using genetic algorithms. In *Proceedings of the IEEE International Parallel and Distributed Processing Symposium*.
17. Zomaya, A. Y., Ward, C., & Macey, B. (1999). Genetic scheduling for parallel processor systems: comparative studies and performance issues. *IEEE Transactions on Parallel and Distributed Systems*, 10(8), 795–812.
18. Lin, M., & Yang, L.T. (1999) Hybrid genetic algorithms for scheduling partially ordered tasks in a multi-processor environment. In *Proceedings of the 6th IEEE Conference on Real-Time Computer Systems and Applications* (pp. 382–387).
19. Woo, S.-H., Yang, S.-B., Kim, S.-D., Han, T.-D. (1997). Task scheduling in distributed computing systems with a genetic algorithm. In *High-Performance Computing on the Information Superhighway, HPC-Asia '97* (p. 301).
20. Hou, E. S. H., Ansari, N., & Ren, H. (1994). A genetic algorithm for multiprocessor scheduling. *IEEE Transactions on Parallel and Distributed Systems*, 5(2), 113–120.
21. Salleh, S., & Zomaya, A. Y. (1999). Scheduling in parallel computing systems: Fuzzy and annealing techniques. Kluwer Academic.
22. Nanda, A., et al. (1992). Scheduling directed task graphs on multiprocessors using simulated annealing. In *Proceedings of the International Conference on Distributed Systems* (pp. 20–27).
23. Boger, M. (2001). *Java in distributed systems*. Wiley.
24. Kunz T. (1991). The influence of different workload descriptions on a heuristic load balancing scheme. *IEEE Transactions on Software Engineering*, 17(7).
25. Yorozu, Y., Hirano, M., Oka, K., & Tagawa, Y. (1987). Electron spectroscopy studies on magneto-optical media and plastic substrate interfaces (Translation Journals style). *IEEE Transaction Journal on Magnetics in Japan*, 2, 740–774.
26. Shivaratri, N. G., Kreuger, P., & Singhal, M. (1992). Load distributing for locally distributed systems. *Computer*, 25(12), 33–44 (reprinted here).
27. Moore, M. (2003). An accurate and efficient parallel genetic algorithm to schedule tasks on a cluster. In *Proceedings of the IEEE International Parallel and Distributed Processing Symposium*.

# Self-driving Cars: An Overview of Various Autonomous Driving Systems



V. Shreyas, Skanda N. Bharadwaj, S. Srinidhi, K. U. Ankith  
and A. B. Rajendra

**Abstract** A car that can navigate by itself without being dependent on human for inputs is known as a self-driving car. There has been a great advancement in automobile industry which is bringing new technologies every day. There are various types of autonomous cars and they are divided based on their level of automation, which includes level 0 to level 5. Advanced methodologies are used to build these cars, and concepts like machine learning and computer vision play a vital role in development of these cars. The accuracy varies based on lots of factors including both internal and external factors. This paper presents survey done on various technologies used in these cars with their results and also about their current trends.

**Keywords** Autonomous car · Self-driving cars · Autonomous vehicle · Neural network

## 1 Introduction

An autonomous car is capable of learning its environment through sensors and camera which will then process the data received through these external devices, which will help in taking decisions. With great development in hardware technology and

---

V. Shreyas (✉) · S. N. Bharadwaj · S. Srinidhi · K. U. Ankith · A. B. Rajendra  
Department of Information Science and Engineering, Vidyavardhaka College of Engineering,  
Mysuru, India  
e-mail: [shreyasvenkatesh06@gmail.com](mailto:shreyasvenkatesh06@gmail.com)

S. N. Bharadwaj  
e-mail: [skanda0017@gmail.com](mailto:skanda0017@gmail.com)

S. Srinidhi  
e-mail: [srinidhis31@gmail.com](mailto:srinidhis31@gmail.com)

K. U. Ankith  
e-mail: [koteraankith@gmail.com](mailto:koteraankith@gmail.com)

A. B. Rajendra  
e-mail: [rabvce@gmail.com](mailto:rabvce@gmail.com)

embedded devices, small components can do computations effectively with all forms of data. There are major car manufacturers including BMW, Google and Tesla that are building and actively testing these cars. Latest results show that autonomous cars have become very efficient and already are driven without any human intervention [1]. Advanced and complicated control systems, algorithms and software take all the sensory data and information to recognize and identify suitable and right routing paths [2, 3]. Autonomous vehicles have complicated control systems that can able to take in sensor data, analyse the data to differentiate different objects in the surrounding environment and identify vehicles and other obstacles in the environment, which will be very helpful for planning to the desired destination [4].

The National Highway Traffic Safety Administration has classified autonomous vehicles as belonging to zero of five levels: i.e. a. no automation, b. assisted automation, c. partial automation, d. high automation and e. full automation. It is important that the driver's attention is needed within level 0 to level 2 modes. Any car that has been manufactured will be considered to be in any one of those levels and also as the level increases, the automation also increases but there are no cars which are in level 5 automation, i.e. high automation, and companies are working towards it. An autonomous car construction implies a mechanical and electrical design. There are two solutions: transform a real car into an autonomous car or to design a new vehicle [5]. Audi expressed that their new A8 would be autonomous up to 60 km/h [6]. The driver does not require safety checks such as frequently gripping the steering wheel. The Audi A8 is claimed to be the first production car to reach level 3 automation, and Audi would be the first car manufacturer to use laser scanners along with cameras and various sensors for their system [7].

Autonomous vehicles have excellent scope as they tend to do fewer errors compared to human drivers. Road accidents are major cause of death in the world and most of these accidents happen due to mechanical problems and driver's distraction. Several perspectives towards vision-based self-driving rely on certain features of the road such as lane markers and systems usually have a specific lane detector [8]. A new study shows that most of the traffic jamming is caused by three major issues: improper car parking, street dwellers and people walking on the roads [9]. These accidents can be brought down if properly trained autonomous vehicles are implemented. Apart from saving lives, consumption of fuel rate can also be reduced with the help of autonomous vehicles. One of the major advantages of autonomous vehicles is that these vehicles can be used in military and this will help in keeping troops out of harm. In 2002, DARPA announced grand challenge which focused on building autonomous vehicles for a prize of \$1 million dollars offered to researchers from top institutes if their vehicles can travel a distance of 142 miles through the Mojave Desert. This was the first grand challenge that took place and later few more challenges took place which motivated researchers to build a self-driving car that can navigate as far as possible.

In this survey, we try to look at and compare various methodologies used in order to build autonomous cars which are popular and also efficient with respect to their performance. There are a good number of concepts and technologies through which autonomous vehicles can be implemented but choosing the right one considering the external factors is prominent. Also, this survey focuses on providing information to the researchers to learn more about trends and technologies in this area since the scope of this area is vast and is open for contributions.

## 2 Background and Related Work

The history of autonomous cars begins in 1920s, where experiments were conducted on self-driving cars. In 1925, radio-controlled car named 'American Wonder' was demonstrated on New York City streets. The trend of radio-controlled cars continued for years, and later in 1980s, vision-guided Mercedes-Benz robotic van was designed in Munich, Germany. Since then, the major focus has been on development of vision-based guiding systems which uses computer vision, LIDAR and GPS technologies. To enhance the research of autonomous system, there are comparisons between various methodologies and their results shown in Table 1.

The concept of neural network training for self-driving cars was introduced in early 1990s by Dean Pomerleau, Carnegie Mellon researcher, how raw images of roads can be captured and trained in order to control the steering of the vehicle based on the condition of roads. Also, this method was considered to be more efficient compared to the other methods used for self-driving cars. This was the major turning point as even to this day self-driving cars use neural network method for training and implementation purposes. There are various neural network training methods like convolutional neural network, R-CNN. The working and performance of neural network vary based on the concepts and more importantly, the input provided. The neural network representation is as shown in Fig. 1.

Neural network training can be done by live streaming of video or using images but the drawback is that it is only limited to 2D images, but LIDAR technology changed everything as it used to gather real-time data of the car surrounding. Spinning light detection and ranging system (LIDAR) was used in order to gather real-time 360-degree maps of the surrounding. LIDAR technology made car even more automated and later cars were using both LIDAR and neural network in order to gather live data and hence made car more autonomous. With the help of built-in GPS, the cars were able to navigate around without any trouble as they had everything to navigate. Also, cars use ultrasonic sensors and odometry sensors for distance measurement and motion detection, respectively. All these components and car controlling system were linked and controlled together by a central computer, which enabled a certain level of automation. Overview of control system is as shown in Fig. 2.



**Table 1** List of various methodologies and their results

Project Name	Year	Methodology	Result
Carnegie Mellon University's Navlab project	1995	Neural networks	98.2% autonomous driving on a 5,000 km cross-country journey [10]
Map free lane following based on low-cost laser scanner for near future autonomous service vehicle	2015	2D laser scanners LMS151, gyro, encoders	The lane detection using 2D based laser is accurate and fast. The path planning based on lane fitting and prediction is reliable and intuitive [11]
Distributed embedded deep learning based real-time video processing	2016	Object detection algorithm, YOLO, DAEDLuS platform	GPU utilization before = 90% Using DAEDLuS platform = 22% [12]
A prototype of an Autonomous police car to reduce fatal accidents in Dubai	2017	R-CNN, kinetic depth sensor, OpenCV	Accuracy of 96% and a mean error of 1.6% with an error of 4% [13]
Autonomous decision making for a driver-less car	2017	CNN, AlexNet, TensorFlow	Final autonomous system is able to provide speed and navigation commands which allows the car to drive in TORCS without any collision or off track driving [14]
Robust lane recognition for autonomous driving	2017	Viola-Jones object detection, AdaBoost training, Hardwareinloop simulation	The designed system for autonomous vehicle guidance can be successfully used to control the driving simulator [15]
Obstacle detection and classification using deep learning	2017	R-CNN, region proposal network, LIDAR	Faster R-CNN was used to build a vision-based object detection system for on-road obstacles [16]
Detecting unexpected obstacles for self-driving cars: fusing deep learning and geometric modelling	2017	Appearance-based semantic detection, stereo-based geometric detection, fusing appearance- and stereo-based detection	Detection rate of over 90% for distances of up to 50 m [17]
End-to-end ego lane estimation based on sequential transfer learning for self-driving cars	2017	CNN, KITTI	Post-processing errors are less when using end-to-to end by which it reduces the re-design and re-optimization [18]

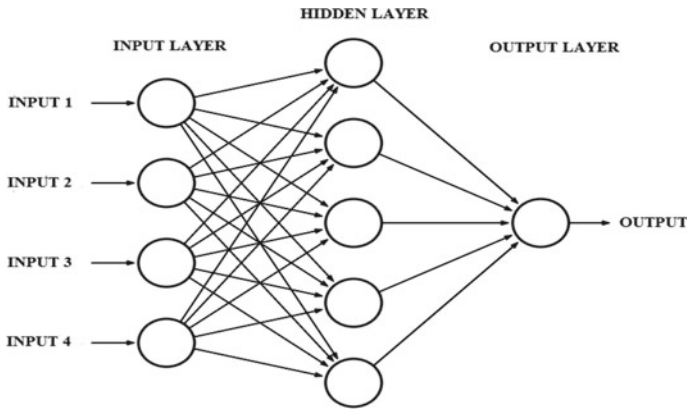


Fig. 1 Neural network representation

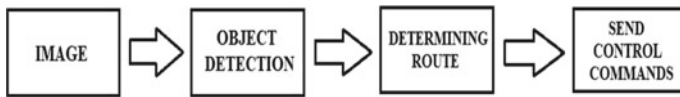


Fig. 2 Overview of control system

### 2.1 Convolutional Neural Network

When we consider a neural network, the network is made up of neurons with weights and biases. CNN is also a neural network, where the neuron takes in several inputs and calculates a weighted sum of the inputs. This is passed through an activation function which generates the output. The difference between neural network and CNN lies in the input given-in neural networks the input given is vector, whereas in CNN the input given is a multi-channelled image. Image recognition, image classifications, object detection, face recognition and so on are some of the major application fields of CNN.

### 2.2 Regional-Based Convolutional Neural Network

In R-CNN of CNN, it is mainly focused on the single region so that the interference is minimized as it expects only the single object of which the interest lies will dominate in the given region. By the method of selective search algorithm, we can detect the regions in the R-CNN, and it is done by resizing the regions with equal size so that it can be used in the CNN classifier and also in bounding box regression. Bounding box regression is needed because the starting proposal might not coincide with the region that is given by the features of CNN.

### **2.3 LIDAR Technology**

Light detection and ranging, or LIDAR, is a remote sensing method which uses light in the form of a pulsed laser. The major components for such a device include a laser, scanner and a GPS receiver. A sensor in LIDAR is used to continuously fire beams of laser light and then determine the time taken for the light to return to the sensor.

### **2.4 AlexNet**

Alex Krizhevsky designed a convolutional neural network called AlexNet, which addresses the problem of image classification. He proposed to take as input an image from one of 1000 various classes, which produces an output that is a vector of 1000 numbers. The element at the  $i$ th position in the output vector is considered as the probability of input image present in the  $i$ th class. As a result, the sum of all elements of output vector is 1. The input image should be an RGB image of size  $256 \times 256$ , which includes all images in training set and testing set as well. If it is not of that size, then it has to be converted to the size before being able to use it for training the network.

### **2.5 AdaBoost**

AdaBoost is an ensemble classifier; ensemble classifiers are made up of multiple classifier algorithms whose output is the combined result of output of those classifier algorithms. AdaBoost classifier forms a strong classifier algorithm by combining various weak classifier algorithms. Using a single algorithm may result in objects being classified poorly. Combining several classifiers by considering the right amount of weight for the selected training set for every iteration can result in greater accuracy for the overall classifier.

### **2.6 TensorFlow**

TensorFlow is an open-source framework that has an artificial intelligence library. Models are built using the data flow graphs generated by this library. Creating a large-scale neural network with many layers is easy for the developer to build. The main application area of TensorFlow includes classification, perception, understanding, discovering, prediction and creation.

## 2.7 Viola–Jones Object Detection

Viola–Jones algorithm is mainly used for the purpose of object detection. The main property of this algorithm is that the detection fast despite the fact that training is slow. Haar basis feature filters are used in this algorithm, so that multiplication is not used.

## 2.8 Hardware-in-the-Loop Simulation

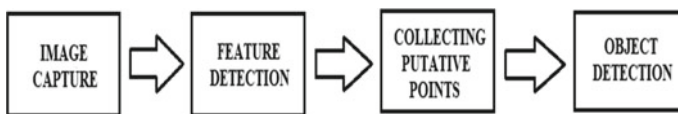
Hardware-in-the-loop simulation or HIL simulation is a type of real-time simulation used for testing control systems. In simple words, the physical part of a machine or system is replaced by a simulation. This provides a rigorous testing method with great variety of benefits. Actuators and sensors are used to connect the machine with the control system.

## 2.9 DEDLUS Platform

It consists of three module: image capture module, data management node and data processing node. The data management node stores the information passed by image capture module capture by webcam or camera; it meagre the image with the result archive by the data processing module and stores it. An object detection algorithm is run on the GPU by the processing module to process the data.

## 2.10 Yolo

It is an object detection technique in which it divides the image into different sections using neural network and predicts the boundaries of each section and then assumes the weight to different sections and identifies the different objects in the image. Steps for object detection are depicted in Fig. 3.



**Fig. 3** Steps in object detection

## ***2.11 High Dynamic Range Imaging***

A greater dynamic range of exposures in images can be obtained in HDR Imaging when compared to the standard methods [19]. The objective here is to present a similar range of luminance as perceived by humans through their eyes. When compared to the traditional methods, HDR images represent a greater luminance in scenarios of real world containing very bright, direct sunlight to extreme shade. A number of different narrower range, exposures of same subject matter is combined after the image is captured. The counterpart of HDR, which is the non-HDR cameras capture images with a limited exposure range, which results in loss of detail in highlights or shadows.

## ***2.12 Deep Learning***

Deep learning is a part of artificial intelligence concerned with surpassing the learning approach used by human beings to gain certain types of knowledge. At a much simpler level, deep learning can be considered as a way of automating the process of predictive analytics. Deep-learning algorithms are arranged in a hierarchy of abstraction and increasing complexity contrary to the traditional machine learning algorithms. The algorithm in the hierarchy makes use of a nonlinear transformation on its input and creates as output a statistical model from what it learned. Once the output has reached an acceptable level of accuracy the iteration is stopped.

## ***2.13 Hebbian Learning Algorithm***

Hebbian learning algorithm tries to explain synaptic flexibility, i.e. about how the synaptic strengths are adapted in brain. If the neurons on the either side of the synapse have linked output, then the synapse linking two neurons is strengthened. When an input neuron triggers, the output neuron triggers, and the synapse is strengthened. By following the resemblance to an artificial system, the weight of tap is increased with high linking between the two sequential neurons.

## **3 Approaches and Working**

Artificial neural network is inspired by the biological neural network, and ANN itself is not the algorithm but it is the combination of many different machine learning algorithms working together to process the given input. This type of system learns to perform task by considering previous example without being programmed to perform

a specific task. It is the combination of connected nodes called artificial neurons. The node receives inputs from an external source or from other nodes and computes the output. Every node has a weight associated with them and with the combination of input value and weight, the node produces an output. This output can be an input to other nodes or can be an output to a system. A typical ANN is the combination of many layers. Different layers perform different kinds of functions on input. Data traverse through the first layer or input layer to the final layer or output layer by traversing through multiple processing layers.

An autonomous car requires five essential functions, i.e., localization, insight, scheduling, vehicle control and system management in order to autonomously drive without human involvement [20]. The system consists of three input nodes camera module, LIDAR, sensors and the combination of these input nodes is called input layer. The camera module takes an image or a video as input, the LIDAR creates 360-degree map of the surrounding of the car while there are many other sensors which sense the environment around the car. This process uses the ANN and machine learning algorithms for processing the data in the hidden layer. Combination of the three inputs will be given as the input to the hidden layer through input node and in each layer many nodes will receive the input from the above layer and will process the input and pass the output to the next layer as its input. The output data from the hidden layer are not part of the global output but are only present inside the network. Based on the weighted combination of its inputs, every single one of these hidden units computes a single real-valued output. The data processed by the hidden layer is given as input to the output layer nodes. The output layer nodes are steering, breaks and acceleration. The steering will control the direction in which the car must move, while accelerator will provide the speed in which the car must travel and the break will be applied when it is needed to control the speed of the car or to stop. The diagram to the above explanation is as shown in Fig. 4.

The system accepts live data from the inputs present in Fig. 4 and will process them to produce output but there are situations where one of these input devices may fail to gather data which will only result in failure of the system. This is how the above system might fail while working. Another possibility is that the system may fail to process and return the output in time and the delay may sometimes lead to an accident. There are many other ways which will result in failure but training system well with all the possible scenarios will result in a strong working system.

## 4 Conclusion

In this paper, we discuss all the existing and current technologies that are used in autonomous vehicle development. Autonomous vehicles have a history starting from early 1920s, and still, there is a huge room for development. Initially, it started with radio-controlled cars, and now, we have cars at level 4, i.e. high automation, and we do not need much time to witness level 5 automated cars. The development of Mercedes-Benz vision-based autonomous van by Ernst Dickmanns and introduction

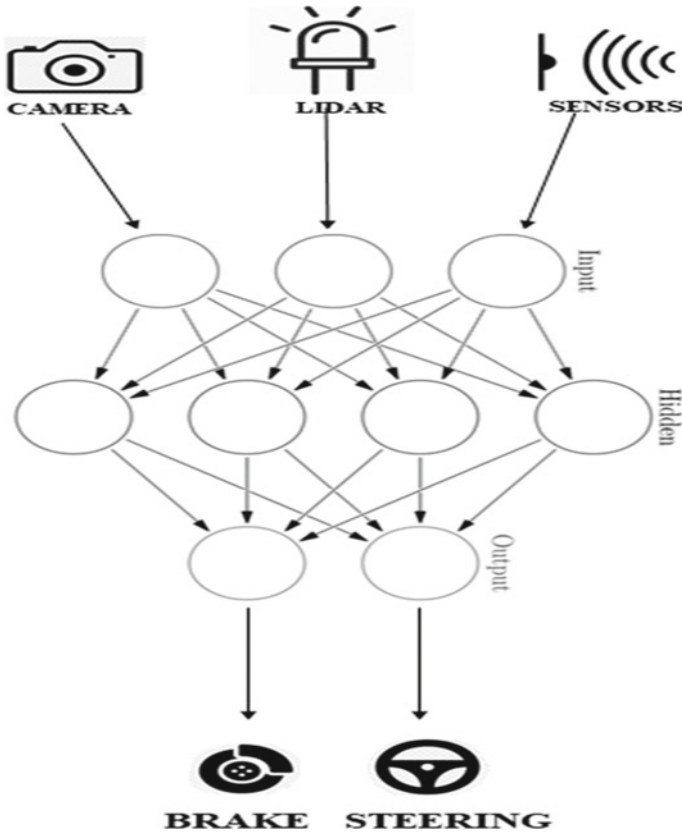


Fig. 4 Working of autonomous car

of neural network in computers Dean Pomerleau in 90s changed everything. Later on, every car that claimed to be an autonomous had vision-based system in them but it had its own limitation and was later equipped with LIDAR technology which made cars more autonomous. Along with those technologies, various types of sensors were used to gather more data, which would make system perform more accurate. Even after usage of all these technologies and equipment, we are still finding hard time to reach level 5 automation because there are certain external factors that cannot be controlled, and those factors do count at times. This shows that still there is a vast scope for new methodologies and development.

## References

1. Tian, Y., & Pei, K. (2018). Automated testing of deep-neural-network-driven autonomous cars. In *Published in 2018 ACM/IEEE 40th International Conference on Software Engineering*.
2. Vishnukumar, H. J., Dr. Müller, C., et al. (2017). Artificial intelligence core for lab and real-world test and validation for ADAS and autonomous vehicles. In *Published in Intelligent Systems Conference 2017*.
3. Liden, D. (2013). What is a driverless car? Wise Geek. 11 Oct 2013.
4. Wentao, Z., et al. (2014). Vehicle detection in driving simulation using extreme learning machine. *Neurocomputing*, 128, 160–165. <https://doi.org/10.1016/j.neucom.2013.05.052>.
5. Pozna, C. (2016). Issues about autonomous cars. In *11th IEEE International Symposium on Applied Computational Intelligence and Informatics*.
6. McAleer, M. (2017). Audi's self-driving a8: Drivers can watch YouTube or check emails at 60 km/h. 11 July 2017.
7. Fayjie, A. R., Hossain, S., et al. (2018). Driverless car: Autonomous driving using deep reinforcement learning. In *Published in 2018 15th International Conference on Ubiquitous Robots*.
8. O' fja'ill, K., Felsberg, M., & Robinson, A. (2016). Visual autonomous road following by symbiotic online learning. IISBN 978-1-5090-1821-5/16/\$31.00.
9. Shamsher, R., & Abdullah, M. N. (2015). Traffic congestion in Bangladesh-causes and solutions: a study of Chittagong metropolitan city. *Asian Business Review*, 2(1), 13–18.
10. Bimbraw, K. (2015). Autonomous cars: Past, present and future.
11. Zhiwei, S., et al. (2015) Map free lane following based on low-cost laser scanner for near future autonomous service vehicle. In *Published in 2015 IEEE Intelligent Vehicles Symposium, South Korea*.
12. Zhang, W., et al. (2016). Distributed embedded deep learning based real-time video processing. In *Published in 2016 IEEE International Conference on Systems, Man, and Cybernetics*.
13. AL Suwaidi, M. A., Al Hammadi, F. J., et al. (2018) A prototype of an Autonomous police car to fatal accidents in Dubai.
14. Gallardo, N., Gamez, N., & Rad, P. (2017). Autonomous decision making for a driver-less car.
15. Kalms, L., Rettkowski, J., et al. (2017) Robust lane recognition for autonomous driving. IISBN 978-1-5386-3534-6/17/\$31.00.
16. Prabhakar, G., Kailath, B., et al. (2017). Obstacle detection and classification using deep learning for tracking in high-speed autonomous driving. IISBN 978-1-5090-6255-3/17/\$31.00.
17. Ramos, S., et al. (2017). Detecting unexpected obstacles for self-driving cars. In *Published in 2017 IEEE Intelligent Vehicles Symposium (IV)*.
18. Kim, J., Park, C. (2017). End-to-end ego lane estimation based on sequential transfer learning for self-driving cars. IISBN 2160-7516/17.
19. Wang, J., & Zhou, L. (2018). Traffic light recognition with high dynamic range imaging. IISBN 1524-9050.
20. Jo, K., Jang, H., et al. (2015). Development of autonomous car-Part II. *IEEE Transactions on Industrial Electronics*, 62(8). (Aug 2015).



# Internet of Things: Industry Use Cases (SAP-HCP)



Avaneesh Kumar Vats and Nagsen Wankhede

**Abstract** Bridge between human and artificial intelligence is the app technology, which has capability of connecting every physical object to app such that fusion of physical world and virtual world of software is enabling enterprises and all its stakeholders to deliver more efficiently which in turn promotes effective decision-making and simplicity.

**Keywords** IoT · SAP-HCP

## 1 Introduction

By connecting people, things, business networks, collection of billion end devices, and transmitting information about every transaction you do knowingly, unknowingly, and eventually guiding you with options to improve your future transactions.

As shown below, an estimated 34 billion IoT installations will be storming the market by 2025 (Fig. 1).

Referring to the above source, huge growth is expected by 2025; however, like any innovation, Internet of Things will have to follow the same path of evolution like any other technological disruptions. IoT solutions available today had been developed with specific challenges in mind, and hence face challenges of interoperability. Various technologies and standards have been applied to build these solutions and consequently they appear as isolated solutions.

The IoT umbrella encompassing all technologies will only be successful when standardized process principles are followed by all stakeholders of the food chain of IoT (UniS), [1].

The article aims to identify the main stakeholders and then their area of operation from a component and industry perspective for an end-to-end implementation of

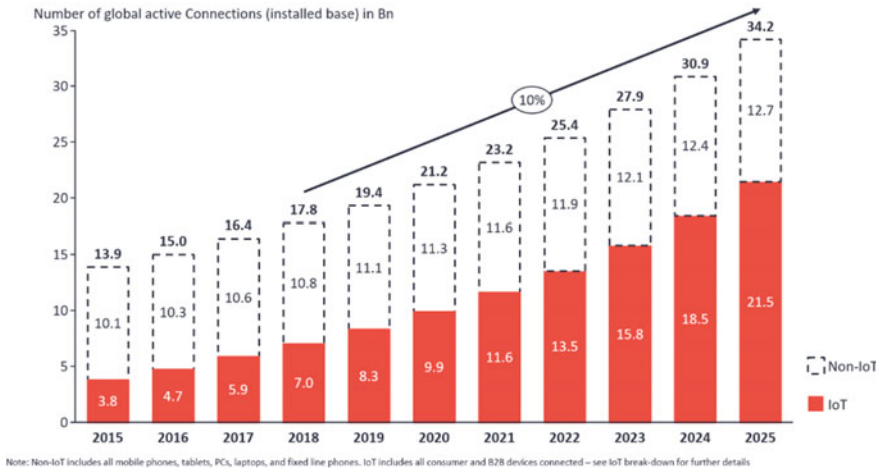
---

A. K. Vats (✉) · N. Wankhede  
Energy Efficiency Services Ltd, EESL, Delhi, India  
e-mail: [avats@eesl.co.in](mailto:avats@eesl.co.in)

N. Wankhede  
e-mail: [nwankhede@eesl.co.in](mailto:nwankhede@eesl.co.in)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_35](https://doi.org/10.1007/978-981-15-0694-9_35)

### Total number of active device connections worldwide



**Fig. 1** Estimated IoT devices by 2025. *Source* <https://iot-analytics.com/state-of-the-iot-update-q1-q2-2018-number-of-iot-devices-now-7b/>

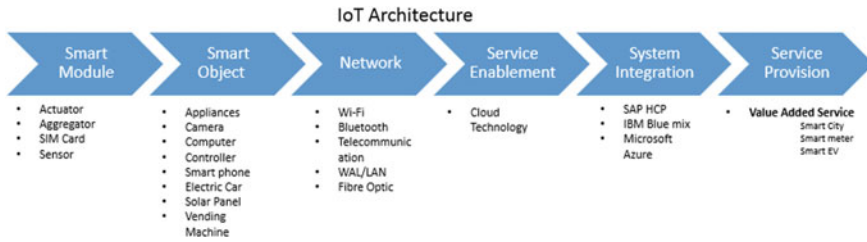
Internet of Things services offered to an end user. It is important to emphasize at this stage that some cool features of a smart device or availability of robust APIs with great features alone cannot build opportunities for IoTs in the marketplace. The combination of all features leading to a differentiated value-added service is what customers are looking for, and hence evaluation of industry-specific use cases is another area of interest to be explored in this article.

## 2 Architecture and Process Stakeholders

The enablement of IoT services is not confined to the features available in individual components but can be achieved by efficient directional data flow through all processes linked with the process chain that refer to the main the component in figure below (Fig. 2).

A brief definition including main features and standards used together with industry stakeholders are provided below:

**Smart device** is a fusion of smart module and smart object. It has the capability of representing an object of physical world to the digital world. Referring to Architecture Reference Model (ARM) this is also known as augmented entity which combines physical entity to a virtual entity containing software resources and services to transmit information to the Internet Marina Ruggieri [2]. A smart device is thus capable of communicating with other devices of Internet.



**Fig. 2** High-level architecture of IoT considering stakeholder involved

In order to categorize as a physical device, it should have the following components (Davy) [3]:

1. Power component—Source of power provided to the device, e.g., battery, solar, electricity, etc.
2. Memory component—Internal memory is in the form of cache, volatile, and nonvolatile to store information for intelligent processing. Processing component—Required for intelligent decision-making.
3. Communication interface—Capability to communicate with other devices through several protocols used, e.g., Ethernet, USB, Bluetooth, GPRS, GSM, etc.

The key enablers of smart devices are as follows:

**Miniaturization:** Remarkable progress made by Microelectromechanical Systems (MEMS) in recent years has expedited the capability of IoT usage.

**Energy requirement:** To power the sensor nodes, several energy harvesting technologies are evolving as battery power is not the most reliable source of energy. Temperature, light, and vibration are the main sources of ambient energy provider to power the sensors.

**Access network technologies:** The range of sensor networks varies from few hundred meters to several miles and is also commonly known as “the last mile”. With the growth of smart sensor capabilities, the backbone becomes a bottleneck and reliance on wireless technologies has increased. Several industry-based standards are used to cover IoT requirements, e.g., Bluetooth 4.0, IEEE 802.15.4e, WLAN IEEE 802.11.

**Network:** 6LoWPAN low-power wireless technology based on IPv6 has emerged to meet IoT requirement.

### 3 Industry Requirement

#### Value-added service

SAP had adopted this to support their customers in the journey of IoT revolution without disrupting the existing business process to offer the value-added services as mentioned below.

#### [A] SAP predictive maintenance

Services can analyze large volumes of operational data and apply predictive insights in real time to increase asset availability and satisfaction levels.

Stakeholders	Value-added services offered
Asset owners	<ul style="list-style-type: none"> <li>• To help asset management</li> <li>• To plan asset outages</li> </ul>
OEMs	<ul style="list-style-type: none"> <li>• To manage warranties</li> <li>• To manage stocks</li> <li>• Optimize enterprise asset management and customer service management</li> </ul>
Dealers and after-sales service	<ul style="list-style-type: none"> <li>• Optimize service delivery</li> <li>• Maximize machine uptime</li> <li>• Increase service efficiency</li> <li>• Reduce cost</li> </ul>

#### [B] SAP connected logistics

It helps the logistics hub to monitor inbound traffic at hub so that it can facilitate communication between all the stakeholders and integrate to backend transportation management systems. Mobile app facilitates complete hands-free operation like handheld scanners.

#### Main features include

- Increased goods throughput and optimized infrastructure for maximum productivity.
- Automate processes for a huge boost in efficiency.
- Shrink waiting times and the need for manual monitoring.
- Reduce emissions and environmental impact.

## 4 Industry Use Cases

### SAP connected manufacturing

An example of the power of smart manufacturing was seen when Harley Davidson worked with SAP to create a factory of the future, which reportedly reduced their production time from 21 days to 6 h. Data collected from connected smart devices has been analyzed on application level to boost the operational efficiency which also predicts the downtime/outages.

### SAP connected logistics:

SAP demonstrates the capability of connected logistics is Hamburg Port Authority (HPA). Expansion of space due to increased demand of containers from 9 to 25 million is addressed here by collaborating with SAP and T-Systems to build a logistics business network.

Leveraging “Telematic” of T-Systems by interfacing telematics data to SAP logistics applications, truck drivers receive real-time information about incidents and parking conditions on mobile devices and can plan routes accordingly. The interconnected logistics network also improves route planning and reduces transportation costs.

## 5 Platform Evaluation

SAP offers a comprehensive solution portfolio using HANA Cloud Platform to build IoT solutions [4].

Features of the platform:

1. Processing very high volume of data at real time using in-memory technology.
2. Ability to process machine, device, sensor, and actuator generated data using SAP\*SQL Anywhere with a lightweight database for end devices.
3. Ability to merge IoT-sensor-generated data to business transactions. Services also include remote device, message, and API management.
4. Enriching decision-making ability by providing predictive and analytical capability. This includes the following capabilities:
  - Text analysis,
  - Geospatial processing,
  - Operational intelligence,
  - Series data processing,
  - Graph engine modeling
  - Multitenant architecture, and
5. Offers flexibility to extend with third party with Open Standards.

## 6 Conclusion

IoT is mainstream now, and more than simply providing real-time connections to objects or devices, IoT is building into a force of astounding change sweeping across the business landscape, knocking down the barriers that separated businesses from their customers, from their employees, and from one another. As no other technology has done before, IoT is opening up vast opportunities for innovation.

SAP HANA Cloud Platform for IoT covers a comprehensive IoT solution portfolio. However, there is still much work to be done. Valuable data needs to be identified and integrated, and partners and customers need to be brought on board within connected networks. The benefits have to be obvious to all.

## References

1. (UniS), F. C. (2013). Internet of Things- Architecture. 25.
2. Marina Ruggieri, H. N. (2013). Internet of things: Converging technologies for smart environments and integrated ecosystems. <http://www.riverpublishers.com/>.
3. Davy, A. (n.d.). Components of a smart device and smart device interactions.
4. \*SAP. (2015). Reimagining business with SAP HANA CLOUD PLATFORM for the internet of things. SAP.
5. SAP. (2014). Connect transform and reimagine business in a hyper connected future.

# **Web and Informatics**

# Organizational Readiness for Managing Large-Scale Data Storage in Virtualized Server Environments



Said Ally

**Abstract** Storage management is a challenging issue due to exponential growth of data and computing workloads transacted per day. Although storage virtualization plays a prominent role in addressing this problem, there exist complexities in the adoption process of virtualization technologies because the transition from a non-virtualized to a virtualized environment is not always a smooth process. Using a survey of 24 public and private institutions in Tanzania, the organizational readiness for managing data storage in virtualized servers is studied. While there is a satisfactory awareness level for most adopters of the existing virtual-based data storage technologies and disk partition techniques, the attention given to securing adopter's practices for virtual resource allocation is inefficient for countering virtual machine attacks. Adopters are prone to starve their virtual machines as soon as their computing workloads expand. The study reveals a serious ad hoc allocation of virtual HDDs with high discrepancy between adopters from the same sector and with similar IT infrastructure, and computing level and demands. Organizations are prompted about the limitations of their current practices which would hardly bring maximum benefits of storage virtualization when expanding from small- to large-scale data workloads.

**Keywords** Storage · Server · Virtualization · Virtual machine · Awareness · Tanzania

## 1 Introduction

Server virtualization is one of the emerging technologies in which adopters face challenges in its adoption process. Being widely adopted in developing countries [1, 2] and rated as the second-most important emerging technology to help achieve IT cost reductions [3–5], very little is known about the adoption of this technology. Although virtualization process has been proven to provide significant IT cost reduction [6, 7],

---

S. Ally (✉)  
The Open University of Tanzania, Dar es Salaam, Tanzania  
e-mail: [said.ally@out.ac.tz](mailto:said.ally@out.ac.tz); [saidallymasomaso@gmail.com](mailto:saidallymasomaso@gmail.com)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_36](https://doi.org/10.1007/978-981-15-0694-9_36)

381



with several other benefits (such as increased availability, scalability, hardware utilization, security, load balancing, isolation [8], and improved system reliability and reduced hardware vendor lock-in [9]), it is clear that organizations are prone to fail to attain the intended economic benefits. One of the major problems in the adoption of server virtualization technologies is organizational readiness to manage the virtualized systems. The proper management of virtualized servers requires skilled staff with security expertise in data storage technologies. Currently, it is not clear which skill components of server virtualization play prominent roles that adopters should focus on. Also, it is important that a critical assessment of adopter's practices toward allocation of storage resources is done to assess organizational readiness for proper management of server virtualization.

Thus, in this paper, we discuss the organizational readiness for managing data storage in virtualized servers with focus on skill level and practices toward virtual resource allocation. The remainder of the paper is organized as follows. In Sects. 2 and 3, critical review of related literature and research methodology is presented, respectively. In Sect. 4, findings and discussions for data storage technology skills and practices are presented. A concluding remark of this work is given in Sect. 5.

## 2 Background

Server virtualization as any other emerging technologies has proven potential impacts [10] on IT asset optimization [4] due to its ability to consolidate, manage, and efficiently distribute server resources [11]. However, one of the major challenges that adopters encounter when incorporating an emerging technology is lack of relevant skills [10, 12]. Data storage technologies and techniques have been considered as the *must-have* skills required for efficient implementation of server virtualization project [13, 14], and hence important sub-themes were derived from these aspects.

### 2.1 Review of Data Storage Technologies

In virtualized server computing, data storage can be implemented using local or shared storage. The choice of local or shared storage has significant impact on the performance of virtual machines and is a key determinant on applicability of some virtualization services such as live migration and load balancing. A local storage is normally located within a virtualized server while a shared storage is implemented using a centralized single storage pool which is accessed by multiple nodes.

Data center architecture for local data storage includes hard disk drives (HDD) and solid-state drives (SSD) [15, 16] coupled by an intelligent processing unit such as a multicore graphic processing unit (GPU) [17]. In terms of performance, the SSDs offer better access performance than HDDs [18]. Basically, SSDs are array of

semiconductor memory available in static mode and are nonvolatile permanent data storage in server side [18].

Storage devices in data centers can be connected using parallel or serial advanced technology attachments (PATA/SATA) disk drive interfaces, network-based iSCSI storage, and serial attached SCSI (SAS) interface storage [19, 20]. Unlike PATA which is widely used at PC level, SATA is mostly used at server level [21].

Data storage in virtualization is also viewed by the way data is accessed through network [18]. Storage systems on network can be implemented using network attached storage (NAS), direct attached storage (DAS), and storage area network (SAN). DAS storage connects directly to a single host server or a cluster of servers instead of directly attached to a network and uses SATA, SAS, external-SATA, and fiber channel (FC) [18, 21]. SAN operates behind servers to provide a common path among storage devices and servers for facilitation of block-level communication between storage and server using a dedicated, high-performance FC channel network. It composes storage device, interconnection network infrastructure (switches), and servers connected to the network, but it is not accessible by other devices through the local area network since it uses separate network. NAS serves files by its hardware, software, or configuration.

## ***2.2 Review of Disk Partition Techniques***

Disk partition technique allows the precise separation or partitioning of a physical HDD into more logical disks for efficient data organization. The logical disks are isolated from each other to increase security. For instance, a BIOS/boot partition is separated from the regular root file system so that the GNU GRUB can use it to load an OS when a GUID partition table (GPT) is contained in a real boot device.

Considering a finite life span of HDDs [18], the stored data are susceptible to inaccessibility caused by media faults unless the HDDs are grouped as array disks using RAID technologies. Use of RAID ensures data redundancy and enhanced performance through making clones of data into two or more disks implemented at different schemes and RAID levels (0, 1, 5, and 10). However, despite its great advantages of data backup and reliability at lower cost in an event of system crash, RAID technology cannot protect the complete data, and it rarely recovers data easily [18], and hence adopters need to be keen throughout implementation of technology.

Therefore, in this paper, an overview of adopter's understanding of data storage technologies and disk partition techniques is presented to explore the organizational readiness for managing data storage in virtualized servers.

**Table 1** Experience and educational level of respondents

Type of adopter	Average years of experience		No of respondents for each education level		
	Since employed	Since virtualized	M.Sc.	B.Sc.	Dip.
Public	8	3	11	20	0
Private	10	6	2	4	1

### 3 Proposed Approach

#### 3.1 Instruments and Population Samples

The focus of the study is based on nine data storage technologies (DST1 to DST9) and four disk partition techniques (DPT1 to DPT4) extracted deductively from relevant sources including local-based shared storage using HDD and SSD [22–25] local-based parallel and serial ATA disk drive interface (PATA and SATA) [25, 26], network-based iSCSI and serial attached storage (SAS) [27, 28], and the cloud-based DAS, NAS, and SAN storage [29, 30].

Data were gathered purposively between June and September 2018 using questionnaires distributed to 38 server admins in 15 public and 9 private entities which have invested in virtualization systems [31].

The awareness level was rated from 1 to 5 scales indicating not at all aware to extremely aware [32]. Responses were collected from various sector-based institutions which have virtualized their servers including banks, telecoms, universities, and government agencies.

#### 3.2 Respondents Demographic Data

Table 1 shows the demographic profiles of the respondents with respect to adopter's category, experience in using VMs, and level of education.

The average level of experience of server admins in private adopters was significantly higher than in public institutions in using systems virtualization. Most server admins in the institutions had bachelor's degrees (24) followed by 13 with master's degrees.

#### 3.3 Reliability Tests

Reliability tests for DST and DPT items calculated using Cronbach's Alpha were found to be 0.809 for  $N = 9$  and 0.847 for  $N = 4$ , respectively, for all 38 server admins. Being greater than 0.7, both values are acceptable Cronbach's Alpha values [33].

## 4 Results and Discussions

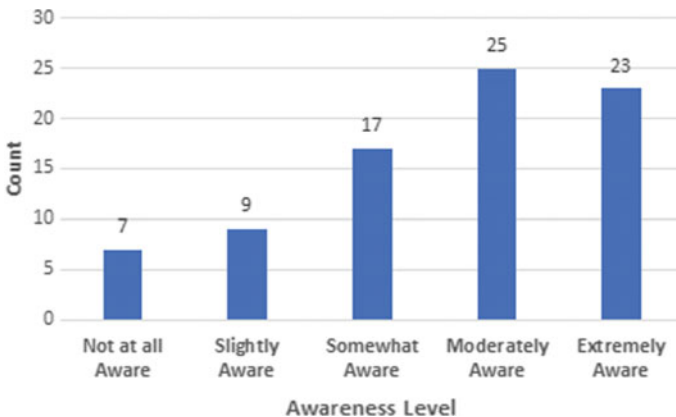
### 4.1 Awareness of Data Storage Technologies

Our Likert scale ordinal data was compared using medians as indicated in Table 2.

As shown in Table 2, all items have median of four except DST1 which has median of five. The results indicate that the DST aspects are basically known by both public and private entities in Tanzania. The overall awareness level for the nine DST aspects is shown in Fig. 1.

**Table 2** Data storage technologies

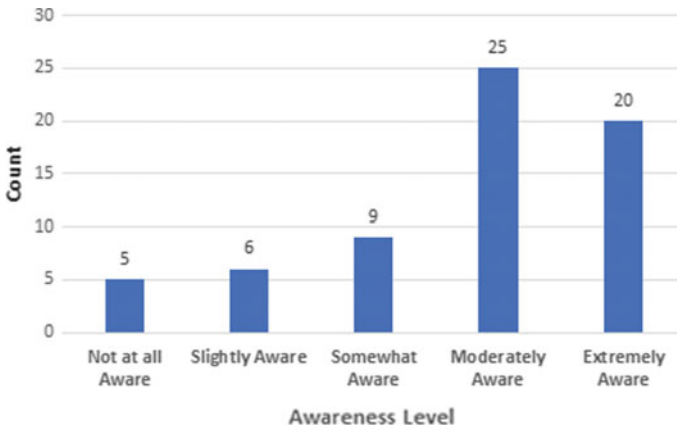
CODE	Awareness level	
	Measured item	Median (N = 38)
DST1	Local-based HDD shared storage	5
DST2	Local-based SSD shared storage	4
DST3	Local-based parallel ATA (PATA) disk drive interface	4
DST4	Local-based serial ATA (SATA) disk drive interface	4
DST5	Network-based iSCSI storage	4
DST6	Network-based serial attached storage (SAS) interface	4
DST7	Cloud-based DAS storage	4
DST8	Cloud-based NAS storage	4
DST9	Cloud0based SAN storage	4



**Fig. 1** Awareness level of data storage technologies

**Table 3** Disk partition techniques

CODE	Awareness level	
	Measured item	Median (N = 38)
DPT1	BIOS partition table	4
DPT2	GUID partition table (GPT)	4
DPT3	Unified extensible firmware interface (UEFI)	4
DPT4	RAID levels	5



**Fig. 2** Awareness level of disk partition techniques

### 4.2 Awareness of Disk Partitions Techniques

As shown in Table 3, similar to DST aspects, the awareness level of DPT techniques was found to be acceptable with all items having median of four except DPT5 which has a median of five.

Figure 2 shows the overall awareness level for the four DPT techniques collected.

### 4.3 Resource Allocation of Virtual Storage

Allocation of virtualized server resources for data storage needs to be managed efficiently to counter security attack because when resource allocation is not controlled, VMs can easily be starved due to scarcity of computing resources.

It was found from the study that the resource allocation for data storage in virtual computing is not dynamically managed due to limitation on the use of VM calculators and resource allocation algorithms such as FCFS, SJF, and UPGT by most adopters. For instance, there is a high discrepancy in the allocation of virtual HDD allocated

per VM with ranges from minimum storage size of 13 GB to maximum size of 64 TB for the adopters of similar IT infrastructure and computing workloads. Discrepancy was also noted when comparing average mean size for the allocated vHDD which is 7,065 GB in public and 2,038 GB in private sectors. Highest values of vHDD were found in finance (banks) and telecoms (mobile) companies as shown in Table 4.

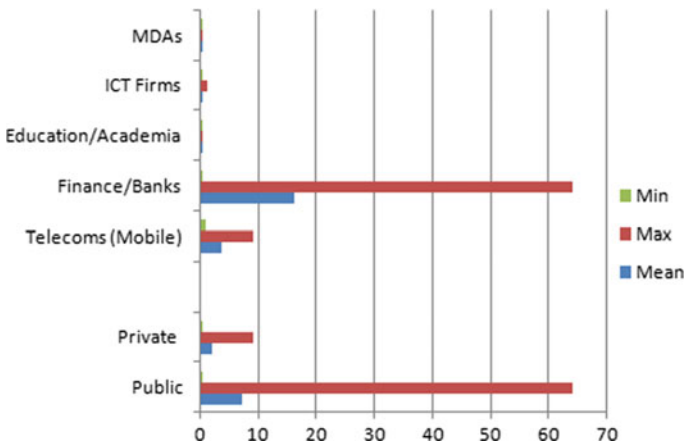
The average mean of 16,085 GB and 362 GB for banks and telecoms, respectively, is another high discrepancy found when allocating vHDDs for adopters of similar infrastructure and computing workloads.

Figure 3 shows vHDD allocation for all adopters.

Furthermore, adopters have set maximum permissible virtual HDD that can be allocated to any VM as a control mechanism for computing resources. Table 5 and Fig. 4 show maximum permissible vHDD size. However, an interesting observation from this aspect is the inconsistency between the size of the allocated vHDDs and the maximum set permissible vHDDs size for adopters of similar services, computing level, and workloads such as banks and telecoms. Although the average mean of

**Table 4** Allocated virtual HDD (in GB) per VM

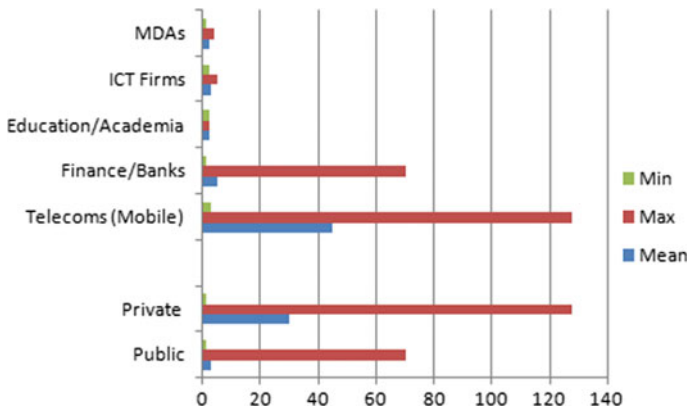
SN	VT adopter	Min	Max	Mean
1	Public	13	64000	7065
2	Private	20	9000	2038
1	Telecoms (Mobile)	700	9000	362
2	Finance/Banks	25	64000	16085
3	Education/Academia	30	30	30
4	ICT firms	13	1000	314
5	MDAs	20	400	257



**Fig. 3** Virtual HDD allocation per VM (in TB) among adopters

**Table 5** Maximum permissible V-HDD size (in TB)

SN	VT adopter	Min	Max	Mean
1	Public	1	20	3
2	Private	1	128	30
1	Telecoms (mobile)	3	128	45
2	Finance/Banks	1	70	5
3	Education/Academia	2	2	2
4	ICT firms	2	5	3
5	MDAs	1	4	2



**Fig. 4** Maximum permissible virtual HDD (in TB) among adopters

vHDDs allocation is considerably much higher in banks than in telecoms, the maximum set permissible vHDDs size for telecoms leads by far when compared to banks despite having similar computing infrastructure, services, and workloads. This is a notable ad hoc practice toward storage management.

The differences in the allocation of virtual HDDs among adopters were measured using statistical computation of a two-way ANOVA as shown in Tables 6 and 7.

As shown in Table 7, the size of the allocated virtual HDD measured against organization type, service organization, and their combination was found to be not significant. This means that there is randomness in the allocation of virtual HDDs. A clear interpretation from this result is the existence of ad hoc practices in the allocation of virtual storage resources.

**Table 6** Allocated V-HDD (in GB) per VM by different adopters

SN	Public adopters	Mean	Std. Dev	N
1	Finance/Banks	32108.00	36826.016	4
2	Education/Academia	30.00	.	1
3	Telecoms (mobile)	700.00	0.000	3
4	MDAs	256.67	126.754	6
5	ICT firms	427.40	523.094	5
	<b>Total</b>	<b>7065.21</b>	<b>20066.125</b>	<b>19</b>
	<i>Private adopters</i>	<i>Mean</i>	<i>Std. Dev</i>	<i>N</i>
1	Finance/Banks	61.75	47.444	4
2	Telecoms (mobile)	8000.00	1414.214	2
3	ICT Firms	30.00	14.142	2
	<b>Total</b>	<b>2038.38</b>	<b>3718.376</b>	<b>8</b>
	<i>Public/private adopters</i>	<i>Mean</i>	<i>Std. Dev</i>	<i>N</i>
1	Finance/Banks	16084.88	29574.114	8
2	Education/Academia	30.00	.	1
3	Telecoms (mobile)	3620.00	4060.419	5
4	MDAs	256.67	126.754	6
5	ICT firms	313.86	469.099	7
	<b>Total</b>	<b>5575.78</b>	<b>16969.122</b>	<b>27</b>

**Table 7** Allocatedv-HDD (in GB): two-way ANOVA

Source	DF	Mean square	F	Sig.
Corrected model	7	487868598.198	2.277	0.073
Intercept	1	239160133.364	1.116	0.304
<b>OrgType</b>	<b>1</b>	310919574.749	<b>1.451</b>	<b>0.243</b>
<b>OrgService</b>	<b>4</b>	358824506.886	<b>1.674</b>	<b>0.197</b>
<b>OrgType*OrgService</b>	<b>2</b>	721489720.121	<b>3.367</b>	<b>0.056</b>
Error	19	214297271.857		
Total	27			
Corrected total	26			

R Squared = 0.456 (Adjusted R Squared = 0.256)



### 4.4 Comparing DST and DPT Awareness for Public and Private Entities

The mean ranks that resulted from a Mann–Whitney U test were used to assess the relationship between the DST and DPT awareness level for both public and private adopters as summarized in Tables 8 and 9.

As shown in Tables 8 and 9, it was found that private adopters have much higher mean ranks than public adopters for all DST and DPT items with exceptions of DST7 and DST9 which represents cloud-based DAS and SAN storages, respectively. The DAS storage is equally known by both public and private. On the other hand, public entities are far more aware of SAN storage compared to private adopters.

However, the awareness differences of DST and DPT aspects between public and private adopters are statistically not significant considering a sample size of  $N = 38$ . When the independent variable (adopter’s type) and the dependent variable (awareness level) are subjected for testing using a Mann–Whitney U test as shown in Tables 10 and 11, all p-values were found to be greater than 0.05. The results of the test are shown in Tables 10 and 11.

**Table 8** Mean ranks for data storage technologies

Item	Adopter’s mean ranks	
	Public (31)	Private (7)
DST1	18.53	23.79
DST2	18.50	23.93
DST3	18.24	25.07
DST4	17.85	26.79
DST5	18.61	23.43
DST6	17.87	26.71
DST7	19.50	19.50
DST8	19.34	20.21
DST9	19.68	18.71

**Table 9** Mean ranks for disk partitions aspects

Item	Adopter’s mean ranks	
	Public (31)	Private (7)
DPT1	18.47	24.07
DPT2	18.16	25.43
DPT3	17.97	26.29
DPT4	18.90	22.14

**Table 10** Mann–Whitney test for data storage technologies

Item	Testing association		
	Mann–Whitney U	Asymp. Sig.	Exact Sig.
DST1	78.500	0.199	0.265
DST2	77.500	0.220	0.249
DST3	69.500	0.127	0.145
DST4	57.500	0.043	0.053
DST5	81.000	0.267	0.317
DST6	58.000	0.046	0.059
DST7	108.500	1.000	1.000
DST8	103.500	0.844	0.854
DST9	103.000	0.827	0.854

**Table 11** Mann–Whitney test for disk partition technologies

Item	Testing association		
	Mann–Whitney U	Asymp. Sig.	Exact Sig.
DPT1	76.500	0.206	0.234
DPT2	67.500	0.107	0.125
DPT3	61.000	0.063	0.076
DPT4	90.000	0.437	0.506

### 4.5 Discussion of the Results

On average, the results indicate a moderate awareness level for the data storage technologies and disk partition techniques when implementing server virtualization. However, this awareness level does not reflect their practices for managing virtual data storage. For example, one-third of adopters cannot realize some virtualization services such as live migration and load balancing due to their dependency on local storage.

On the other hand, adopters whose data were stored on shared storage using SAN and NAS devices were found to face difficulty in selecting compatible network devices and efficient connection method that match their computing requirements, thus affecting server performance due to increased traffic.

Another key interpretation was that adopters currently feel safe because they have small computing workloads, but with current exponential data growth in virtualized environment, adopters are susceptible to virtualization sprawl that may easily leave most of the virtual machines starved due to scarcity of storage space. Additionally, their moderate awareness level does not match their skill level, and hence a need for adopters to build in-house capacity remains vital.

## 5 Conclusions

This paper has explored organizational readiness for managing large-scale data storage in virtualized environments. With current practices, adopters will find it difficult to accommodate the exponential growth of data that need to be stored and computed considering the rapid advancement of data storage technologies and techniques. Adopters need to invest a lot in relevant required expertise for storage virtualization to attain a secured and cost-effective virtualized computing infrastructure. To attain a reliable performance for VMs and consequently circumvent a chance of one VM to control and consume all the storage space while leaving other VMs starving, there is a need for the adopters to have clear operational guidelines on the resource allocation with use of virtualization calculators for efficient estimation of the required resources while considering hardware specification.

This work can be extended to assess compatibility of storage virtualization techniques for each of the widely used hypervisors.

**Acknowledgements** This study aimed at investigating the organizational readiness for managing large-scale data storage in virtualized server environments. A survey of 38 server administrators from 15 public and 9 private institutions from banks, telecoms, universities, and government agencies participated in the study and given consent for the study done. Due to exponential growth of data and computing workloads transacted per day, participants found the study worth mentioning in addressing the complexities during the transition process from a non-virtualized to a virtualized environment with a use of storage virtualization. Lastly, for a smooth conduct of this research, it is honor to convey special acknowledgement to the management of The Open University of Tanzania for financing this research.

## References

1. Tsai, P. (2016). Server virtualization and OS trends. Network Article, Spiceworks. Retrieved April 14, 2017, from <https://community.spiceworks.com/networking/articles/2462-server-virtualization-and-os-trends>.
2. Malla, M. (2017). The top 5 trends in cloud and virtualization testing for 2017. Spirent predictions. Spiceworks. Retrieved June 24, 2017, from <http://vmblog.com/archive/2017/01/09/spirent-2017-predictions-the-top-5-trends-in-cloud-and-virtualization-testing-for-2017.aspx#.Wo64WzNRWiM>.
3. Ogunyemi, A. A., & Johnston, A. K. (2012). Exploring the roles of people, governance and technology in organizational readiness for emerging technologies. *The African Journal of Information Systems*, 4(3), Article 2. <https://digitalcommons.kennesaw.edu/ajis/vol4/iss3/2>.
4. Padhy, R. P. (2012). Virtualization techniques and technologies: state-of-the-art. *International Journal of Global Research in Computer Science*, (UGC Approved Journal), 2(12), 29–43.
5. Kanoongo, B., Jagani, P., Mehta, P., & Kurup, L. (2014). Exposition of solutions to hypervisor vulnerabilities. *International Journal of Current Engineering and Technology*, 4(5), 3244–3247. International Press Corporation (INPRESSCO), ISSN 2277–4106.
6. Uhlig, R., Neiger, G., Rodgers, D., Santoni, A. L., Martins, F. C., Anderson, A. V., et al. (2005). Intel virtualization technology. *Computer*, 38(5), 48–56. <http://dx.doi.org/10.1109/mc.2005.163>.

7. Infographic. (2014). Top 5 reasons why you can't avoid not to virtualize. *Infographic whitepaper on vmware*. Retrieved July 11, 2017, from <https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/infographic/vmw-top5-reasons-infographic.pdf>.
8. HP. (2009) Server virtualization technologies for x86-based HP blade system and HP ProLiant server's technology brief (3rd ed.) Hewlett Packard Development Company, L.P.
9. Chiramal, H. D., Mukhedkar, P., & Vettathu, A. (2016). *Mastering KVM virtualization* (1st ed), Birmingham, UK: PACKT Publishing Ltd. ISBN 978-1-78439-905-4.
10. Fleischer, T., Decker, M., & Fiedeler, U. (2005). Assessing emerging technologies-methodological challenges and the case of nanotechnologies. *Technological Forecasting and Social Change*, 72, 1112–1121.
11. Ally, S. (2018). *A holistic approach to determine security levels of virtual machines in the adoption and use of open source hypervisors*. Ph.D. thesis, The Open University of Tanzania, pp. 320.
12. Cetindamar, D., Phaal, R., & Probert, D. (2009). Understanding technology management as a dynamic capability: A framework for technology management activities. *Technovation*, 29, 237–246.
13. Osakiand, N., Yamamoto, A. (2009). Method and apparatus incorporating virtualization for data storage and protection. Hitachi Ltd. U.S. Patent 7,594,072.
14. Scarfone, K. (2011). *Guide to security for full virtualization technologies*. DIANE Publishing.
15. Yang, Z., Bhimani, J., Wang, J., Evans, D., & Mi, N. (2017). Automatic and scalable data replication manager in distributed computation and storage infrastructure of cyber-physical systems. *Journal of Scalable Computing*, 18(4). Special Issue on Communication, Computing, and Networking in Cyber-Physical Systems.
16. Maskalik, S., Aravind, S., Debashis, B., Sachin, T., & Allwyn, S. (2017). Multi-spoke connectivity of private data centers to the cloud. U.S. Patent Application 14/981,424, filed March 2, 2017.
17. Yang, Q. (2018). Storage of data reference blocks and deltas in different storage devices. Western Digital Technologies Inc, 2018, U.S. Patent Application 10/108,348.
18. Patil, P. T. (2016). A study on evolution of storage infrastructure. *International Journal*, 6(7).
19. Rezaei, A., Suri, T., & Brennan, B. (2018). Versioning storage devices and methods. Samsung Electronics Co Ltd, 2018. U.S. Patent Application 10/061,523.
20. Dube, S. J., Ahmed, S., Shetty, S. V., & Palmer, J. R. (2017). System and method for providing management network communication and control in a data center. Dell Products LP, 2017. U.S. Patent Application 15/015,961.
21. Choi, Y. S., & Kwon, H. J. (2017) SATA host bus adapter using optical signal and method for connecting SATA storage using the same. Electronics and Telecommunications Research Institute, 2017. U.S. Patent 9,722,702.
22. Wu, W., Wei, X., Zibin, Y., & Qingbin, L. (2018). Exploring the potential of coupled array of SSD and HDD for multi-tenant. In *2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, pp. 653–657. IEEE.
23. Bernal, E. R., & Heninger, I. M. (2019). Addressing usage of shared ssd resources in volatile and unpredictable operating environments. January 3, 2019.
24. Chang, H. P., Yu-Cheng, Y., & Chung, P. Y. (2018). Design and implementation of a shared multi-tiered storage system. In *2018 3rd International Conference on Computer and Communication Systems (ICCCS)*, pp. 94–98. IEEE.
25. Yashwanth, Y. K. (2018). Different forensic tools on a single SSD and HDD, their differences and drawbacks.
26. Yoo, B., Seongjin, L., & Youjip, W. (2018). Analyzing the storage defects from the perspective of synthetic fault injection. *Journal of Information Science and Engineering*, 34(1), 1–20.
27. Aazam, M., Eui-Nam, H., & Marc, S. (2018). Towards media inter-cloud standardization—evaluating impact of cloud storage heterogeneity. *Journal of Grid Computing*, 16(3), 425–443.
28. Raj, P., & Anupama, R. (2018). Software-defined storage (SDS) for storage virtualization. In *Software-defined cloud centers* (pp. 35–64). Springer, Cham.

29. Park, J. K., & Jaeho, K. (2018). Big data storage configuration and performance evaluation utilizing NDAS storage systems. *AKCE International Journal of Graphs and Combinatorics*, *15*(2), 197–201.
30. Lanka, A., & Ariun, G. (2018). Remotely accessible, low power network attached storage device. In *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, pp. 1083–1088. IEEE.
31. Silverman, D. (2010). *Doing qualitative research—A practical handbook* (3rd ed.). SAGE Publications Ltd. ISBN: 978-1-84860-033-1.
32. Vagias, W. M. (2006). Likert-type scale response anchors. Clemson International Institute for Tourism & Research Development, Department of Parks, Recreation and Tourism Management, Clemson University.
33. DeVellis, R. F. (2003). *Scale development: Theory and applications—Applied social research methods* (3rd ed.). Thousand Oaks, Calif: Sage Publications. ISBN-13: 978-1412980449.

# Classification of Forest Cover Type Using Random Forests Algorithm



Arvind Kumar and Nishant Sinha

**Abstract** Natural resource planning is an important aspect for any society. Knowing forest cover type is one of them. Multiple statistical and machine learning approaches are already proposed in past for classification. In the current work, publicly available dataset of Forest Cover type (FC) from UCI repository was taken and classified for the forest cover type, using random forests machine learning algorithm. On ten-fold cross validation, we got accuracy of 94.6% over 70.8% accuracy of original work presented at UCI repository. The result is also compared with existing work done in past and it have been shown that random forests algorithm performed better than most of existing works.

**Keywords** Classification · Random forests · Machine learning · Unbalanced data · Forest cover type

## 1 Introduction

To develop strategies for ecosystem, we require descriptive knowledge of available forest lands. This descriptive information facilitates easy and perfect decision-making process for resource planners [1]. This type of information may be either collected by doing fieldwork manually or by doing satellite image processing, remote sensing, Geographic Information System (GIS), predictive modeling, etc. The fieldwork is done by human and cost intensive. Whereas predictive modeling techniques are easy to implement and are of low cost. Again, these predictive modeling may be based on either statistical modeling or through Machine Learning (ML) techniques [2]. In recent years, researchers have given more attention to generalized approach of

---

A. Kumar (✉)  
Bennett University, TechZone II, Greater Noida 201310, India  
e-mail: [arvind.jki@gmail.com](mailto:arvind.jki@gmail.com)

N. Sinha  
Pitney Bowes Software, C28-29, Sector-62, Noida 201301, India  
e-mail: [nishant.sinha@pb.com](mailto:nishant.sinha@pb.com)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_37](https://doi.org/10.1007/978-981-15-0694-9_37)

ensemble machine learning based classification. In ensemble approaches, multiple classifiers are designed and trained and based on voting process, their results are combined. Random Forests (RF) are an ensemble approach of decision tree-based classifier that utilize a improved method of bootstrapping as bagging [3, 4].

Blackard et al. [1, 5], studied and introduced Forest Cover type (FC) dataset, which is now publicly available at University of California (UCI) Knowledge Discovery in Databases (KDD) Archive [6]. They analyzed this dataset first time using a feedforward Artificial Neural Network (ANN) model and Linear Discriminant Analysis (LDA) method based statistical model. Their artificial neural network model got a classification accuracy of 70.58% and linear discriminant analysis model got an accuracy of 58.38%. Multiple other researches also studied this dataset and used different techniques for classification and prediction of forest cover type. A summary of various research is presented in Table 4. On the same FC dataset, we use a machine learning algorithm, named random forests algorithm, and got classification accuracy of 94.59% in this work.

Section 2 of this paper gives details about current work available in literature and Sect. 3 gives details about dataset and formalizes the problem statement. Section 4 introduced the random forests algorithm used in this research work and implementation details. Section 5 gives results and Sect. 6 conclude this research work.

## 2 Current Work

In this work, for classification of forest cover type, we have used dataset available on UCI website. As per data details, this dataset contains four different types of forest lands which are situated in the Roosevelt National Forest of northern Colorado, United States [6]. In this proposed work, over 12 cartographic measurement of different parameters summarized in Table 1 have been used as independent variable. There are seven major forest land cover types found in the mentioned area. These seven classes are considered as dependent variables in proposed work. In original work [5], two classification approaches were presented by authors

1. A feedforward artificial neural network model.
2. A Linear Discriminant Analysis (LDA) method based statistical model.

Their ANN model got a classification accuracy (70.58%) and linear discriminant analysis model got an accuracy of 58.38%. There are multiple approaches proposed by researchers to classify the FC type. Table 4 gives a brief of available work done in the past. This contains authors, their algorithmic approaches, and best found classification accuracy in their work.

**Table 1** Data attribute information

Attribute name	Measurement unit	Attribute description
Elevation	Meters	Elevation
Aspect	Azimuth	Aspect in degrees
Slope	Degrees	Slope
Horizontal distance to hydrology	Meters	Horizontal distance to nearest surface water features
Vertical distance to hydrology	Meters	Vertical distance to nearest surface water features
Horizontal distance to roadways	Meters	Horizontal distance to nearest roadway
Hill shade at 9 am	0–255 (Index)	Hill shade index at 9am, summer solstice
Hill shade at noon	0–255 (index)	Hill shade index at noon, summer solstice
Hill shade at 3 pm	0–255 (Index)	Hill shade index at 3pm, summer solstice
Horizontal distance to fire points	Meters	Horizontal distance to nearest wildfire ignition points
Wilderness area (4 binary columns)	0 (Absence) 1 (Presence)	Wilderness area designation
Soil type (40 binary columns)	0 (Absence) 1 (Presence)	Type of soil
Forest cover type (7 types)	1–7	Type of forest cover

### 3 Dataset Description

The dataset used in this work is available on UCI KDD archive [7], which is taken from United State Forest Service (USFS) and United State Geological Survey (USGS) data. Four wilderness forest areas found in Roosevelt National Forest of northern Colorado, United States, have been taken into consideration here. Total 12 cartographic attributes are taken as independent variables and forest cover type is taken as dependent variable. Table 1 gives a brief detail about this. The dataset contains total of 581012 observation points, with 36.46, 48.76, 6.15, 0.47, 1.63, 2.99, and 3.53% data of class 1, 2, 3, 4, 5, 6, and 7, respectively. Table 2 contains details about record count. Thus, the dataset has unbalanced class distribution. Unbalanced data classification has its own classification complexity. For example, a classifier having good accuracy may give best prediction for majority classes but worst prediction for minority classes. As data point for minority classes are low, accuracy will be still high. Therefore, a good classifier should have a balance between classification accuracy over all classes, i.e., minority as well as majority classes [3].



**Table 2** Forest cover type class distribution

	Forest type	Record count	Record percentage
1	Spruce Fir	211840	36.46
2	Lodgepole Pine	283301	48.76
3	Ponderosa Pine	35754	6.15
4	Cottonwood/Willow	2747	0.47
5	Aspen	9493	1.63
6	Douglas-fir	17367	2.99
7	Krummholz	20510	3.53
	Total	581012	100.00

## 4 Random Forests

Random forests (RF) is an ensemble technique used for classification. The term came from random decision forests that was first proposed by Tin Kam Ho of Bell Labs in 1995 [8]. In RF, we grow multiple trees as opposed to a single tree in decision tree model. Thus, RF is an ensemble classifier that contains multiple random Decision Trees (DT). Individual DTs classification output is taken and these values are combined to generate final output of classifier. Random forests considers a small subset of features for splitting at each classification and regression tree-like classifiers [9]. In summary, RFs utilized the principle of stochastic modeling to construct a tree-based classifier. In randomly selected subspace of feature space, multiple classification decision trees are built. For a  $m$ -dimensional feature space, there are  $2^m$  possible subspace, in which a decision tree may be created. The use of stochastic approach is a flexible way to explore different available possibilities.

---

### Algorithm 1 Random forests algorithm

---

- 1: From training set take  $K$ -data points randomly.
  - 2: Construct decision trees associated to above  $K$ -data points.
  - 3: Choose the number  $N_{tree}$  of trees you want to build and repeat step 1 and 2.
  - 4: For any new data point,
    - a. Predict the value of  $Y$  using each of  $N_{tree}$  trees for the data point in consideration.
    - b. Assign a new data point across all the predicted  $Y$  values.
-

## 5 Results

For predicting forest cover type, we designed a classifier based on random forests algorithm. For its implementation, python PANDAS library is used for data processing and SKLEARN is used for algorithm implementation. To evaluate the performance of proposed work, quality measurement metrics like precision, recall, F-1 score, and accuracy are calculated [10]. There were total 581012 data points. For performance comparison, ten-fold cross validation is done. Please refer to Table 3 for classification summary. We got an accuracy of 94.6% over 70.8% accuracy of original work. Table 4 compares this result with other works done in past and available in literature. This is clear from this comparison that our results are better in terms of classification accuracy and other metrics.

**Table 3** Result summary

Class	F1-score	Precision	Recall
1	0.95	0.94	0.95
2	0.95	0.95	0.96
3	0.94	0.93	0.95
4	0.87	0.90	0.83
5	0.82	0.93	0.74
6	0.89	0.92	0.86
7	0.95	0.97	0.93

**Table 4** Forest cover type classification performance comparison

References	Accuracy	Approaches
De Almeida et al. [11]	80.50	Dynse K-E92 method
Krawczyk and Wozniak [12]	75.76	One-class classification model with different incremental learning and forgetting procedures
Rojanavasvu et al. [13]	72.90	sUpervised Classifier System (UCS)
Narayan et al. [14]	64.32	k-dimensional tree with Repeated Bisection technique
Prudhomme and Lallich [15]	67.80	Kohonen Opt technique with normalization of the attributes
Garcke and Griebel [16]	70.30	sparse grid
Castro et al. [17]	76.00	Fuzzy ARTMAP (FAM) with data partitioning in regions
Oza [18]	77.87	Bagging with multilayer perceptrons

(continued)

**Table 4** (continued)

References	Accuracy	Approaches
Secretan et al. [7]	79.28	No Matchtracking Fuzzy ARTMAP
Pal and Mitra [19]	67.01	Rough-fuzzy algorithm
Liu et al. [20]	88.00	k-NN classification with IOC algorithm
Koggalage and Halgamuge [21]	89.72	Support vector machine classifier, classification for class 1, 2 and 5 only
Liu et al. [22]	90.00	Decision tree classifier
Mitra et al. [23]	67.75	k-NN Algorithm
Yang and Webb [24]	68.60	Naïve Bayes with nondisjoint discretization method
Demiriz et al. [25]	74.75	AdaBoost
Frank et al. [26]	82.50	Loglikelihood pruning
Furnkranz [27]	66.80	R3
Lazarevic and Obradovic [28]	71.00	Progressive sampling technique with a boosting algorithm
Lazarevic and Obradovic [29]	73.20	Distributed boosting technique
Kaburlasos and Petridis [30]	70.00	Backpropagation
Blackard [1]	70.58	Artificial neural network
<b>Proposed Work</b>	<b>94.60</b>	Random forests algorithm

## 6 Conclusion

In this paper, Forest Cover type (FC) data from UCI Knowledge Discovery in Databases Archive is taken to train and predict forest cover type. To design classifier, random forests machine learning algorithm is used. The complete suite is implemented in Python. To evaluate proposed work performance, we calculated different quality measurement metrics. On ten-fold cross validation, we got an accuracy of 94.6% over 70.8% accuracy of original work. In Table 4, result is also compared with other works done in past, which show that RF performed better than most of existing works.

**Acknowledgements** The authors would like to thank Mr. Tejalal Choudhary for the helpful discussions and support during this work.

## References

1. Blackard, J. A., & Dean, D. J. (1999). Comparative accuracies of artificial neural networks and discriminant analysis in predicting forest cover types from cartographic variables. *Computers and Electronics in Agriculture*, 24(3).
2. Badulescu, L. (2017). Data mining classification experiments with decision trees over the forest covertype database. In *21st International Conference on System Theory Control and Computing (ICSTCC)*, 236.
3. Xue, J.-H., & Hall, P. (2015). Why does rebalancing class-unbalanced data improve auc for linear discriminant analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(5), 1109–1112.
4. Mentch, L., & Hooker, G. (2016). Quantifying uncertainty in random forests via confidence intervals and hypothesis tests. *The Journal of Machine Learning Research*, 17(1), 841–881.
5. Blackard, J. A. (2000). Comparison of neural networks and discriminant analysis in predicting forest cover types. Ph.D. Dissertation, Department of Forest Sciences, Colorado State University, Fort Collins, Colorado.
6. UCI Machine Learning Repository: Covertype Data Set. <https://archive.ics.uci.edu/ml/datasets/covertype>.
7. Secretan, J., Castro, J., & Georgiopoulos, M. (2005). Parallelizing the fuzzy artmap algorithm on a beowulf cluster. In *Proceedings. 2005 IEEE International Joint Conference on Neural Networks* (Vol. 2, p. 475).
8. Tin Kam Ho. (1995). Random decision forests. *Document Analysis and Recognition*, 1, 278–282.
9. Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2006) Random forests for land cover classification. *Pattern Recognition Letters*, 27(4).
10. Lever, J., Krzywinski, M., & Altman, N. (2016). Points of significance: Classification evaluation. *Nature Methods*, 13.
11. De Almeida, P. R. L., et al. (2016). Handling concept drifts using dynamic selection of classifiers. In *2016 IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI)*. IEEE.
12. Krawczyk, B., & Woźniak, M. (2015). One-class classifiers with incremental learning and forgetting for data streams with concept drift. *Soft Computing*, 19(12).
13. Rojanavas, P., et al. (2009). A self-organized, distributed, and adaptive rule-based induction system. *IEEE Transactions on Neural Networks*, 20(3), 446–459.
14. Narayan, B. L., Murthy, C. A., & Pal, S. K. (2006). Maxdiff kd-trees for data condensation. *Pattern Recognition Letters*, 27(3), 187.
15. Prudhomme, E., & Lallich, S. (2005). Quality measure based on kohonen maps for supervised learning of large high dimensional data. In *Proceedings of Applied Stochastic Models and Data Analysis (ASMDA'05)* (pp. 246–255).
16. Garcke, J., & Griebel, M. (2005). Semi-supervised learning with sparse grids. In M. R. Amini, O. Chapelle, & R. Ghani (Eds.), *Proceedings of ICML* (pp. 19–28). Workshop on Learning with Partially Classified Training Data.
17. Castro, J., Georgiopoulos, M., Secretan, J., DeMara, R. F., Anagnostopoulos, G., & Gonzalez, A. (2005). Parallelization of fuzzy artmap to improve its convergence speed: The network partitioning approach and the data partitioning approach. *Nonlinear Analysis: Theory Methods & Applications*, 63(5).
18. Oza, N. C. (2005). Online bagging and boosting. In *Proceedings of IEEE International Conference on Systems Man and Cybernetics* (Vol. 3, pp. 2340–2345). New Jersey, USA. Special Session on Ensemble Methods for Extreme Environments.
19. Pal, S. K., & Mitra, P. (2004). Case generation using rough sets with fuzzy representation. *IEEE Transactions on Knowledge and Data Engineering*, 16(3).

20. Liu, T., Yang, K., & Moore, A. W. (2004). The ioc algorithm: Efficient many-class non-parametric classification for high-dimensional data. In *Proceedings of the 2004 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining-KDD'04* (pp. 629–634). New York, USA: ACM Press.
21. Koggalage, R., & Halgamuge, S. (2004). Reducing the number of training samples for fast support vector machine classification. *Neural Information Processing-Letters and Reviews*, 2.
22. Liu, X., Bowyer, K. W., & Hall, L. O. (2004). Decision trees work better than feed-forward back-prop neural nets for a specific class of problems. In *2004 IEEE International Conference on Systems, Man and Cybernetics* (Vol. 6, pp. 5969–5974). Hague, Netherlands.
23. Mitra, P., Murthy, C. A., & Pal, S. K. (2002). Unsupervised feature selection using feature similarity. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3).
24. Yang, Y., & Webb, G. I. (2002). Non-disjoint discretization for naive-bayes classifiers. In A. G. Hoffmann Sammut (Ed.) *Proceedings of the 19th International Conference on Machine Learning (ICML'02)* (pp. 666 – 673). San Francisco, USA: Morgan Kaufmann Publishers Inc.
25. Demiriz, A., Bennett, K. P., & Shawe-Taylor, J. (2002). Linear programming boosting via column generation. *Machine Learning*, 46(1).
26. Frank, E., Holmes, G., Kirkby, R., & Hall, M. (2002). *Racing committees for large datasets* (pp. 153–164). Berlin: Springer.
27. Fürnkranz, J. (2001). Round robin rule learning. In *Proceedings of the 18th International Conference on Machine Learning (ICML-01)* (pp. 146– 153).
28. Lazarevic, A., & Obradovic, Z. (2001). Data reduction using multiple models integration. In *Proceedings of the 5th European Conference on Principles of Data Mining and Knowledge Discovery (PKDD'01), Germany* (pp. 301–313), Berlin: Springer.
29. Lazarevic, A., & Obradovic, Z. (2001). The distributed boosting algorithm. In *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining-kdd'01* (p. 311) San Francisco, USA: ACM Press.
30. Kaburlasos, V. G., & Petridis, V. (2000). Fuzzy lattice neurocomputing (fln) models. *Neural Networks*, 13(10), 1145.

# Impact of Noisy Labels in Learning Techniques: A Survey



Nitika Nigam, Tanima Dutta and Hari Prabhat Gupta

**Abstract** Noisy data is the main issue in classification. The possible sources of noise label can be insufficient availability of information or encoding/communication problems, or data entry error by experts/nonexperts, etc., which can deteriorate the model's performance and accuracy. However, in a real-world dataset, like Flickr, the likelihood of containing the noisy label is high. Initially, few methods such as identification, correcting, and elimination of noisy data was used to enhance the performance. Various machine learning algorithms are used to diminish the noisy environment, but in the recent studies, deep learning models are resolving this issue. In this survey, a brief introduction about the solution for the noisy label is provided.

**Keywords** Noisy labels · Deep learning approach · Non-deep learning approach

## 1 Introduction

Deep learning is one of the latest emerging areas of machine learning which enables computer to learn from the experience [12]. In general, the deep learning model imitates the working of human brain to process the data for decision-making. The term learning refers to be an iterative process which helps to enhance the knowledge gain. So, the capability of good decision-making depends upon the datasets. The dataset is the collection of real-world data which is used in deep learning model. It is further classified into two parts: training dataset and test dataset [12]. The training dataset is used in learning process of model for classification of new samples. Therefore, the noiseless dataset is an essential requirement for training model. But, the analysis of

---

N. Nigam (✉) · T. Dutta · H. P. Gupta  
Department of Computer Science and Engineering, IIT (BHU), Varanasi, India  
e-mail: [nitikanigam.rs.cse18@iitbhu.ac.in](mailto:nitikanigam.rs.cse18@iitbhu.ac.in)

T. Dutta  
e-mail: [tanima.cse@iitbhu.ac.in](mailto:tanima.cse@iitbhu.ac.in)

H. P. Gupta  
e-mail: [hariprabhat.cse@iitbhu.ac.in](mailto:hariprabhat.cse@iitbhu.ac.in)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_38](https://doi.org/10.1007/978-981-15-0694-9_38)

real-world applications always produces a large set of values that will not inevitably have truth values, known as noisy data. The “Noisy data” referred as mislabeled data was discovered in 1980s, a [1, 27] and yet the research is also active in current years [13, 37]. The presence of noise in the dataset obscured the relationship between the features of an object and its class [14]. This will increase the complexity problem in classification of dataset.

Noise is categorized into an attribute (feature) noise and class noise [44]. The modification in observed values are known as feature noise and mislabelling of a class in the observed labels allocated to an instance are known as class noise [44]. In [30], the non-systematic errors are referred to as noise, in which the values of attributes, class attributes, or both have been altered. Some authors [37] categorizes the noise into two parts: symmetric and asymmetric noise. The noise due to uniform distribution are known as symmetric noise and asymmetric noise occur due to fixed-rule flipping [37].

The class noise is said to be a label noise if observed labels are polluted, i.e., incorrectly labeled [9]. The root of label noise involves data entry error by a specialist or annotators, inadequate availability of resources or knowledge provided to tag a label for a particular object, inter-experts advice on the identical article, or low-budget nonexperts are employed for labeling the data [6]. In [9], the communication and encoding of data are reported to be mislabeled noisy data, in the real-world applications, approximately 5% of encoding and communication fault refers to noise [28]. The existence of noise in the dataset leads to degradation of performance [26, 42] in classification models. This paper focuses on label noise which is a subset of class noise. The essential demand is to learn from the accurate data as the noisy data are the mislabeled data which furnish the wrong information. Thus, the accurate prediction depends on the learning of a machine with the trustworthiness of labeled data. But, the probability of getting noiseless data is inadequate in case of real-world applications, e.g., Flickr provides the image dataset but mostly consist of the noisy labeled dataset as the data has been indicated by humans.

The main contribution of this paper are summarized as follows:

- We review about the noisy data and their effect in classification problems.
- This paper also focuses on the various approaches to resolve the aforementioned problem.

The rest of the paper is organized as follows. The next section describes about the noise, its source, and consequence of noisy labels. Section 3 is the related work, classified into a deep learning approach and non-deep learning approach to resolve the problem of classification with corrupted labels. We conclude the paper in Sect. 4.

## 2 Noisy Labels: Definition, Source, and Consequences

Noise is an irregular patterns present in the dataset but is not a part of real data. In [14], noise is defined as the ambiguous relation between the features and its class. The ubiquity of noise in the data may alter the essential characteristic of an object.

Due to this, the performance of classification may degrade [14]. There are challenges due to noise which may affect the intrinsic properties in classification problem which are discussed as follows:

- It may create small clusters of data points of a particular class in the area of domain which belongs to the different class.
- It eliminates instances located in key areas within an appropriate class.
- It also disrupts the boundary of the classes and increases overlapping among them.

The noise is classified into attribute noise and class noise, which are described as follows [38, 44]:

- The class noise occurs due to misclassification and inconsistent examples. The source of class noise also includes contradictory examples and mislabelling of an instance. In general, it occurs on the boundaries of classes where the characteristics similarity between the datapoints is maximum.
- Attribute noise incorporates unavailable values, error in values, and incomplete information. An alteration in the values of attributes of an instance are also referred as attribute noise.

Attribute noise are more harmful than class noise because erroneous attribute values is uncertain and random.

A lot of work had been done in the field of removing the class noise which results in improving the classification accuracy. The difficult task is to remove attribute noise which is still in research. The class labels are intentionally corrupted by an adversary, which are known as noisy labels [3].

The source of noise are also relevant which are discussed as follows [9]:

- Distribution.
- Data entry error.
- Inefficient data description to tag a class of a instance.
- Subjective error.
- Data communication and encoding problem.
- Nonexperts decision-making.
- Instances of overlapping near boundary.
- Biological artifacts.

The presence of noise is a key issue in the real-world dataset and has many drawbacks. The consequences of label noise are given by

- Decrement in the performance of classification [21]. Some authors have shown experimentally that noisy label leads to degradation in performance [26, 42].
- The complexity of learned models is increased, for instance, the increment in the number of nodes of decision trees [9].
- The examined instances of the possible classes can be changed [9].



### 3 Approaches Used to Resolve Noisy Data

The problem of learning with noisy labels is a serious issue in any classification model. There are following sources such as Mechanical Turk or any crowdsourcing platform, i.e., Google and Microsoft for obtaining large dataset, but the possibility of noise in the dataset is very high. Numerous non-deep learning and deep learning approaches have been proposed to resolve this problem in various applications. Here is a concise explanation of the related work below and in Table 1, summary of some methods are shown.

#### 3.1 Non-deep Learning Approach

Some non-deep learning strategies have been proposed to deal with noisy label dataset. These are statistical methods to reduce the noise and enhance the performance.

In 1999, the identification of mislabeled noise was proposed for a small dataset. The author applied ensemble classifiers such as decision tree, linear classifier, and k-nearest neighbor (KNN) in addition with majority and consensus filtering scheme [6, 18, 39]. Majority and non-objection are the two threshold systems for classification of noise. After identification and correction of noisy labels, the uncomplicated approach is to eliminate the noise, here in [45], mislabeled data is first recognized and then eliminated from the large dataset. The error count variables are used to compute the noisy label instances, and the higher probability occurrence tagged as the noisy data. Another method was proposed for correcting the mislabeled data [38] by comparing the polished (clean) dataset with the noisy dataset.

Bagging and boosting are the two strategies for the detection of noisy data in which weights are assigned to the training sample. The weights are updated on each turn which are labeled as noisy labels if the count value exceeded a threshold value [10, 11, 43]. Ada-boosting algorithm is used for identification and detection of noise by comparison with the clean dataset for only binary classes and in [4], robust boosting algorithm is integrated with Adaboost for identification of mislabeled data. A ranking-based method is also used for the detection of noise. An ensemble-based ranking method is recommended in [34], noisy instances are ranked based on the occurrence of inaccurate predictions done by a learner.

Some research also concentrates on the relabeling of noisy data, which helps in eliminating the noisy data from crowdsourcing or any website. In [20], the author proposed a non-deep learning approach that attempts relabeling of noisy labels and on each round, the noisy labels, as well as predictors, are updated. There is some probabilistic modeling based method which was provided by [31] for supervised learning to determine the problem generated by multiple experts data. In this, the noisy labels are removed by majority votes. The same concept is extended by [41] for the multi-annotators problem through supervised and semi-supervised learning.

**Table 1** Summary of methods for noisy labels

Statistical method		Deep learning method	
Surrogate loss	Majority and nonobjective methods [39]	Robust loss function	1. Defense mechanism for various architecture [40]
Bagging and Boosting	Data Cleaning Method [17, 36]		2. Noisy level reweight [19]
	Data tolerant Method [4, 7, 29]	3. Cross-entropy loss [16]	
		4. Global ratio [15]	
Probabilistic method	Cluster based approach [5]	Modeling the latent labels	5. Bootstrap function [32]
Noise rate estimation method	Noise elimination approach [14]		1. Parallel classifier

The additional work done was that a model was proposed to learn from unlabeled data in the semi-supervised environment as well as to handle missing annotators.

Some approaches are based on noise rate estimation which is shown in [24]. The author applied class probability estimator to learn from altered binary labels by utilizing the order-statistics on a predicted array of numbers.

### 3.2 Deep Learning Approach

Different techniques have been proposed to deal with noisy labels in various applications using machine learning models. In recent years, the deep learning model overshadows the previous techniques. Noisy data problem is resolved by using deep learning techniques. It is classified into two methods which is robust loss function and modeling the latent labels. In [2], CNN deep learning model has been used for image classification, in which an auxiliary image regularization procedure is used to depreciate the noisy data. Bootstrapping and importance re-weighting [22, 32] are the techniques to resolve the dilemma of statistical outliers, i.e., random classification noise. The bootstrapping [32] method was given by Yarowsky in 1995, is a self-learning method which helped in handling subjective as well as unlabeled images. Re-weighting is another technique which is used to alleviate the problem to depreciate the noisy labels [22]. The information is acquired from a small clean dataset and with the help of knowledge graph, noisy labels have been updated.

The method to protect from acquiring the noisy labeled dataset by deep neural network, defense mechanisms (adversarial attack and distillation) have been proposed. In [40], the author introduced multiple defense mechanisms to protect the deep

architecture from learning the noisy data by using robust loss function for adversarial attacks besides using cross-entropy softmax loss function. Above literature is based on robust loss function to diminish the effect of noise.

Some work has been done in the field of modeling the latent labels to guide the classifiers accurately. It also helps in establishing a change for adaption from latent labels to noisy labels. The first work was performed for random and symmetric noise in which [25] proposed a latent model for aerial images. This approach was extended for different noise, in [35] used the convolutional neural network to train on the small and large noisy dataset. Two models have been proposed to minimize the noisy data which significantly improves the performance. In [8], two-step strategy is used to train the convolutional neural network. In the first step, the simple web images have been taken from Google image search which is learned by CNN's to grab the knowledge. Further, in the following step, Flickr images are used to train, but it is challenging to train CNN's with noisy data. Thus, the refining method has been done before providing as input to CNN. The refining method is based on the similarity relationship graph; if there is an existence of error, then it is back-propagated through graph so as to get the accurate classification.

Some methods are based on electing the samples to mark noisy labels, for example, in [23], decoupling method is used which maintains two predictors. The updation of predictors is dependent upon disagreement samples. But, this approach is dependent upon selecting sample and bias selected sample may lead to an error. Co-teaching method is given by [13], in which two distinct deep neural networks (DNN) have been trained concurrently, and in every mini batch dataset, they teach each other. The updation policy for a noisy label is the same as [23], i.e., on disagreement noisy data is updated. There is a probabilistic approach to eliminate the noisy data using deep learning, in [33], an algorithm has been stated which cleans the noisy labels on the postulate of majority voting method.

In [37], an unsupervised method is used to determine the best stop time before learning the noisy data. This method is known as limited gradient descent (LGD) which is applied to both symmetric and noisy asymmetric labels. A reverse sample is created from a given dataset which is different from the main pattern; the analysis significantly determines that only large scaled clean patterns are to be learned.

## 4 Conclusion

The presence of noise in data is a common problem that produces several negative consequences in classification problems. This survey summarized that the noisy data is a complex problem and harder to provide an accurate solution. In general, the data of real-world application is the key source of noisy data. There are two approaches to handle noisy labels. In the deep learning approach, different architectures are implemented for the elimination of noisy labels. The method of elimination of noisy labels in deep learning approach is further classified into a robust loss function and modeling latent variable. The statistical-based methods have been discussed in the

non-deep learning approach in which mostly algorithms were based on majority voting mechanism, bagging and boosting method, noise rate estimation, and the probabilistic method.

All these approaches improve the performance but can't eliminate the noisy labels completely. In recent work, limited gradient descent method has been used for getting the best stop time before learning the noisy labels without eliminating the noisy labels. The work on symmetric noise and asymmetric noise has been ignored. In future, the scope of work could be done in the above field.

**Acknowledgements** This work is supported by Science and Engineering Research Board (SERB) file number ECR/2017/002419, project entitled as A Robust Medical Image Forensics System for Smart Healthcare, and scheme Early Career Research Award.

## References

1. Angluin, D., & Laird, P. (1988). Learning from noisy examples. *Machine Learning*, 2(4), 343–370.
2. Azadi, S., Feng, J., Jegelka, S., & Darrell, T. (2015). Auxiliary image regularization for deep cnns with noisy labels. [arXiv:151107069](https://arxiv.org/abs/1511.07069).
3. Biggio, B., Nelson, B., Laskov, P. (2011). Support vector machines under adversarial label noise. In *Asian Conference on Machine Learning* (pp. 97–112).
4. Bootkrajang, J., Kabán, A. (2013). Boosting in the presence of label noise. [arXiv:13096818](https://arxiv.org/abs/1309.6818).
5. Bouveyron, C., & Girard, S. (2009). Robust supervised classification with mixture models: Learning from data with uncertain labels. *Pattern Recognition*, 42(11), 2649–2658.
6. Brodley, C. E., & Friedl, M. A. (1999). Identifying mislabeled training data. *Journal of Artificial Intelligence Research*, 11, 131–167.
7. Cantador, I., Dorronsoro, J. R. (2005). Boosting parallel perceptrons for label noise reduction in classification problems. In *International Work-Conference on the Interplay Between Natural and Artificial Computation* (pp. 586–593). Springer.
8. Chen, X., Gupta, A. (2015). Webly supervised learning of convolutional networks. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 1431–1439).
9. Frénay, B., & Verleysen, M. (2014). Classification in the presence of label noise: A survey. *IEEE Transactions on Neural Networks and Learning Systems*, 25(5), 845–869.
10. Freund, Y., Schapire, R. E., et al. (1996). Experiments with a new boosting algorithm. In *ICML, Citeseer* (Vol. 96, pp. 148–156).
11. Friedman, J., Hastie, T., Tibshirani, R., et al. (2000). Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *The annals of statistics*, 28(2), 337–407.
12. Goodfellow, I., Bengio, Y., Courville, A., Bengio, Y. (2016). *Deep learning* (Vol. 1). MIT press Cambridge.
13. Han, B., Yao, Q., Yu, X., Niu, G., Xu, M., Hu, W., et al. (2018) Co-sampling: Training robust networks for extremely noisy supervision. [arXiv:180406872](https://arxiv.org/abs/1804.06872).
14. Hickey, R. J. (1996). Noise modelling and evaluating learning from examples. *Artificial Intelligence*, 82(1–2), 157–179.
15. Izadinia, H., Russell, B. C., Farhadi, A., Hoffman, M. D., Hertzmann, A. (2015) Deep classifiers from image tags in the wild. In *Proceedings of the 2015 Workshop on Community-Organized Multimodal Mining: Opportunities for Novel Solutions* (pp. 13–18). ACM.
16. Joulin, A., van der Maaten, L., Jabri, A., Vasilache, N. (2016). Learning visual features from large weakly supervised data. In *European Conference on Computer Vision* (pp. 67–84). Springer

17. Karmaker, A., & Kwek, S. (2006). A boosting approach to remove class label noise 1. *International Journal of Hybrid Intelligent Systems*, 3(3), 169–177.
18. Khoshgoftaar, T. M., Zhong, S., & Joshi, V. (2005). Enhancing software quality estimation using ensemble-classifier based noise filtering. *Intelligent Data Analysis*, 9(1), 3–27.
19. Li, Y., Yang, J., Song, Y., Cao, L., Luo, J., & Li, L. J. (2017). Learning from noisy labels with distillation. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 1910–1918)
20. Lin, C. H., Weld, D. S., et al. (2014). To re (label), or not to re (label). In *Second AAAI Conference on Human Computation and Crowdsourcing*
21. Liu, H., & Zhang, S. (2012). Noisy data elimination using mutual k-nearest neighbor for classification mining. *Journal of Systems and Software*, 85(5), 1067–1074.
22. Liu, T., & Tao, D. (2016). Classification with noisy labels by importance reweighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(3), 447–461.
23. Malach, E., Shalev-Shwartz, S. (2017). Decoupling “when to update” from “how to update”. In *Advances in Neural Information Processing Systems* (pp. 960–970).
24. Menon, A., Rooyen, B. V., Ong, C. S., Williamson, B. (2015). Learning from corrupted binary labels via class-probability estimation. In F Bach, D Blei, (Eds.) *Proceedings of the 32nd International Conference on Machine Learning, PMLR, Lille, France, Proceedings of Machine Learning Research* (Vol. 37, pp. 125–134). <http://proceedings.mlr.press/v37/menon15.html>.
25. Mnih, V., Hinton, G. E. (2012). Learning to label aerial images from noisy data. In *Proceedings of the 29th International Conference on Machine Learning (ICML-12)* (pp. 567–574)
26. Nettleton, D. F., Orriols-Puig, A., & Fornells, A. (2010). A study of the effect of different types of noise on the precision of supervised learning techniques. *Artificial Intelligence Review*, 33(4), 275–306.
27. Oja, E. (1980). On the convergence of an associative learning algorithm in the presence of noise. *International Journal of Systems Science*, 11(5), 629–640.
28. Orr, K. (1998). Data quality and systems theory. *Communications of the ACM*, 41(2), 66–71.
29. Oza, N. C. (2004) Aveboost2: Boosting for noisy data. In *International Workshop on Multiple Classifier Systems* (pp. 31–40). Springer.
30. Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1), 81–106.
31. Raykar, V. C., Yu, S., Zhao, L. H., Valadez, G. H., Florin, C., Bogoni, L., et al. (2010). Learning from crowds. *Journal of Machine Learning Research*, 11(Apr), 1297–1322.
32. Reed, S., Lee, H., Anguelov, D., Szegedy, C., Erhan, D., Rabinovich, A. (2014) Training deep neural networks on noisy labels with bootstrapping. [arXiv:14126596](https://arxiv.org/abs/1412.6596).
33. Rodrigues, F., Pereira, F. C. (2018). Deep learning from crowds. In *Thirty-Second AAAI Conference on Artificial Intelligence*.
34. Sluban, B., Gamberger, D., & Lavrač, N. (2014). Ensemble-based noise detection: Noise ranking and visual performance evaluation. *Data Mining and Knowledge Discovery*, 28(2), 265–303.
35. Sukhbaatar, S., Bruna, J., Paluri, M., Bourdev, L., Fergus, R. (2014). Training convolutional networks with noisy labels. [arXiv:14062080](https://arxiv.org/abs/1406.2080).
36. Sun, J. W., Zhao, F. Y., Wang, C. J., Chen, S. F. (2007). Identifying and correcting mislabeled training instances. In *Future Generation Communication and Networking (FGCN 2007)* (Vol. 1, pp. 244–250), IEEE.
37. Sun, Y., Xu, Y., et al. (2018). Limited gradient descent: Learning with noisy labels. [arXiv:181108117](https://arxiv.org/abs/1811.08117).
38. Teng, C. M. (1999). Correcting noisy data. In *ICML, Citeseer* (pp. 239–248)
39. Verbaeten, S., Van Assche, A. (2003). Ensemble methods for noise elimination in classification problems. In *International Workshop on Multiple Classifier Systems* (pp. 317–325). Springer.
40. Vu, T. K., Tran, Q. L. (2018). Robust loss functions: Defense mechanisms for deep architectures. In: *2018 10th International Conference on Knowledge and Systems Engineering (KSE)* (pp. 163–168). IEEE.
41. Yan, Y., Rosales, R., Fung, G., Subramanian, R., & Dy, J. (2014). Learning from multiple annotators with varying expertise. *Machine Learning*, 95(3), 291–327. <https://doi.org/10.1007/s10994-013-5412-1>.

42. Yao, J., Wang, J., Tsang, I. W., Zhang, Y., Sun, J., Zhang, C., et al. (2019). Deep learning from noisy image labels with quality embedding. *IEEE Transactions on Image Processing*, 28(4), 1909–1922.
43. Zhong, S., Tang, W., & Khoshgoftaar, T. M. (2005). *Boosted noise filters for identifying mis-labeled data*. Department of Computer Science and engineering, Florida Atlantic University.
44. Zhu, X., Wu, X. (2004). Class noise vs. attribute noise: A quantitative study. *Artificial Intelligence Review*, 22(3), 177–210.
45. Zhu, X., Wu, X., Chen, Q. (2003). Eliminating class noise in large datasets. In *Proceedings of the 20th International Conference on Machine Learning (ICML-03)* (pp. 920–927)

# Performance Analysis of Schema Design Approaches for Migration from RDBMS to NoSQL Databases



Basant Namdeo and Ugrasen Suman

**Abstract** State-of-the-art database paradigm allows data to generate from different types of devices but these data have no fixed format according to RDBMS structure. Due to industrial need for business expansion, these data need to be restructured into effective database. Therefore, there is a need for migrating existing data into new database technology such as NoSQL which can efficiently handle them. NoSQL is a group of technologies, which does not follow the concept of relational database. NoSQL stores information in different types of data model such as column oriented, document oriented, or graph based. These models have their own way of storing information and schema design. Various research works have been performed in finding appropriate schema design for migration of data from relational model to NoSQL. In this paper, we have performed performance analysis of different database schema design for migration from RDBMS to NoSQL databases. The selected replication schema design provides better performance in execution time.

**Keywords** Database migration · Schema design · NoSQL · RDBMS

## 1 Introduction

Database schema design provides the way to the application and database developer to efficiently interact with database. A proper data model or database schema is important for efficient working of software application. As the data is growing exponentially, new database technology has emerged and business needs have been changed. With increasing growth of data, legacy database systems such as RDBMS have failed to fulfill customer expectations. Data is growing rapidly in each instance of time and users of that data expect reading data as fast as possible. On the other

---

B. Namdeo (✉)

International Institute of Professional Studies, DAVV, Indore, India

e-mail: [basant\\_nd@yahoo.com](mailto:basant_nd@yahoo.com)

U. Suman

School of Computer Science and IT, DAVV, Indore, India

e-mail: [ugrasen123@yahoo.com](mailto:ugrasen123@yahoo.com)

© Springer Nature Singapore Pte Ltd. 2020

M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,

Lecture Notes in Networks and Systems 94,

[https://doi.org/10.1007/978-981-15-0694-9\\_39](https://doi.org/10.1007/978-981-15-0694-9_39)

hand, RDBMS has some restrictions on storing large amount of data in clusters of computer because of their ACID nature. RDBMS follows the relation model in which data is stored in tables, and these tables are linked together by foreign key and referential key. When user wants to read information that is stored in multiple tables, they have to join the tables in SQL.

Joins are the most common part available in SQL for availing information from multiple tables, but joins are not supported in most of the NoSQL databases. Executing joins are slow in nature. But NoSQL technology relaxes ACID properties, and works on BASE (Basically Available, Soft state and Eventually consistence) model [1]. NoSQL stores related information (i.e., linked table data) in one place and these can be retrieved as a whole and can be stored as a whole in one place. As a result, schema design plays an important role in NoSQL to retrieve and store the information efficiently. RDBMS is popularly used in designing effective databases. On the other hand, NoSQL provides flexibility in database design as compared to RDBMS. Existing companies are moving toward NoSQL technology from RDBMS to gain the advantages of NoSQL.

NoSQL is an umbrella term, which is used for a group of non-relational database management systems [2]. Some of the NoSQL software supports SQL or SQL like query languages. It provides good horizontal scalability for storing and retrieving information. NoSQL databases can be broadly categorized into document model, graph model, and columnar data model [3]. In document model, data is stored in a collection of documents. Each document typically uses a BSON (Binary JSON-binary-encoded serialization of JSON-like documents) structure for storing information in document. MongoDB and CouchDB use document model. In graph model, data is stored in graph structure, which has nodes and edges [4]. Neo4j and Giraph use this model. In columnar data model, key-value are grouped in column family. Each column family may have different number of key-value pairs. Software such as Cassandra, hbase, etc. uses columnar model.

In this paper, we mainly focus on performance analysis of different schema design approaches for migration from RDBMS to NoSQL databases. The paper is organized as follows. Section 2 presents the literature survey on various approaches for RDBMS to NoSQL schema design in document, columnar, and graph databases. Section 3 provides various schema design approaches in NoSQL. Experimental analysis is performed for different schema design approaches of NoSQL in Sect. 4. Finally, the conclusion of this paper is presented in Sect. 5.

## 2 Approaches for RDBMS to NoSQL Schema Design

There are three approaches for migration from RDBMS to NoSQL schema design, namely, relational to columnar, relational to document, and relational to graph database. These approaches are discussed in the following subsections.



## 2.1 *Relational to Columnar*

Columnar database uses the concept of wide table, in which at the time of table creation only column families are defined. Later, when data is stored in the table, each column family can have different number of column name–value pairs. In this, each row of table may have different number of column–value pairs in a single column family. The difficulty with RDBMS to NoSQL, especially columnar database, is finding or identifying column families, and which information is stored in which column family.

A workload-driven approach is developed to create schema design for columnar data stores [5, 6]. The NoSE (NoSQL Schema Evaluator) uses conceptual schema of database and workload as SQL queries (which we frequently executed on RDBMS) as input. They have used Binary Integer Program (BIP) method for availing the optimized database schema. In this approach, workload must be given prior to NoSE system, and if workload is not properly estimated then column families will not be created properly. An algorithm is proposed for finding the row-key for a table in columnar database [7, 8]. Row-key is important in table because it uniquely identifies a row in table, and in columnar database it is a combination of all the columns of related tables of RDBMS. In this approach, same table is referenced by more than one table which causes duplication of data. Also, during data updation we have to update same information in more than one places.

Li presented a heuristic-based approach for transforming relational data model to NoSQL, specifically wide column data model [9]. Several rules are formulated for various tasks, such as what should be column family, how foreign keys should be handled, etc., for converting relational table into HBase table [10]. Zhao et al. propose a method for handling multiple nested tables in HBase [11]. However, this handles the multiple nested tables, but it stores nested table’s information in more than one location, and this paper does not address the 1-m relation between multiple nested tables.

Vajk et al. propose an algorithm for denormalizing relational tables, by which related tables can be grouped into one new table. They provide some possible schema of database, and later finds cost-optimal schema among them, which will be imported in column-oriented database Table [12]. Yangua et al. propose some rule for transforming a multidimensional conceptual model into two NoSQL ones, column-oriented and document-oriented models [13]. They have performed a comparative study between both the NoSQL database models.

Above literature survey of relational to columnar database reveals that there are various ways of schema design in columnar database. Main focus of various authors is finding columns of relational database tables, which can be grouped to form the column family.

## 2.2 *Relational to Document Store*

Document-oriented databases store the information in collection of documents. Here, RDBMS table can be viewed as a collection, and row of table is viewed as document. But, as RDBMS table has fixed column design, this model relaxes it. Different documents have different number of key-value pairs. It stores the related information as sub-documents. The problem with RDBMS to NoSQL, especially document database, is finding or identifying which group of information will be stored as sub-document or as a separate document.

An algorithm is proposed for converting relational tables into MongoDB [14]. They have designed some rules in an algorithmic format for converting one-to-one, one-to-many, and many-to-many relation of entity relationship model into document database. They define which tables will be treated as individual document, which will be treated as embedded document, etc. Karnitis et al. provide solution for schema design of document-oriented database [15]. According to their work, first they find the meta-model of RDBMS, and then they apply depth-first search and breadth-first search based algorithms for finding a path from root table to other related tables. This path is used for creating document template of root table. They classify each table in RDBMS into four subclasses, namely, simple entity, complex entity, N:N link, and codifier. According to their subclass, they define table as separate document or sub-document of parent document. This works good, but this algorithm adds too many weakly related tables in root node, or say, in main document, and this also required database expert intervention for refining database schema.

Zhao et al. also proposed a schema conversion model from RDBMS to NoSQL [16]. They store relational schema information as a graph model, and then convert this graph into documents model. They give some rules for converting tables, which have relationship like vertical (A references B, B references C, and so on) and horizontal (A references B1, B2, Bn) extensions. Yoo et al. give solution for storage problem in schema design in NoSQL by column-level denormalization [17]. Generally, when transformation of data from RDBMS to NoSQL is performed, table-level denormalization is considered, but they suggested only non-primary-foreign-key column information will be duplicated in documents of NoSQL. For deriving of NoSQL schema, we must have sql query as input or workload, but later if the design of sql query is changed then this column-level denormalization process must be done again, and so NoSQL database schema will be changed.

Above literature survey of relational to document database reveals that there are various ways of schema design in document database. Main focus of various authors is finding column's or related table's data of relational database tables, which can be grouped to form the sub-documents of main documents.

### 2.3 Relational to Graph Database

Graph database uses the graph model for storing the information. It uses vertex or node for storing data, and edges for making relation between nodes. Edges can also store information. A row of a table of RDBMS becomes the node. Relationship, which is defined by foreign key in relational database, becomes the edges of graph, and table name becomes label name. Each node stores the information by means of key–value pair, and may have different number of key–value pairs for different nodes of same label name. Following paragraph describes various approaches followed by various research groups for converting relational database to graph database.

In a simplest way, relational database can be transformed into graph database by simply converting each tuple of a table into node, grouping set of node by label name just like table name groups set of tuples. Foreign keys between tables are transformed into edges of graph, which are connecting two nodes [18]. Park et al. proposed 3NF equivalent graph 3EG transformation. They used relational database, which is in 3NF as an input, and give four rules for converting relational database table into graph database [19]. By taking 3NF database as input, all the nodes in graph database have only those keys as a property, which are completely dependent on node's primary key. Virgilio et al. also proposed an algorithm for converting relational to graph database [20]. They aggregate key–values of more than one row in one node, so that when user needs related information, they can get it from visiting only one node. They define some rules for various cases for grouping different row's information in one node. They created group of full schema path, which goes from source node to destination node, and then find which node's information will be duplicated in which node.

Above literature survey of relational to graph database reveals that there are various ways of schema design in graph database. Main focus of various authors is to find columns of relational database tables, which can be grouped to form the key–values of single node, and which key–values can be sub-grouped into a single node and finds the relationship of nodes.

## 3 Schema Design Approaches in NoSQL

There are various approaches available for schema design for migration from relational database to NoSQL databases. NoSQL database approaches have different nature with respect to various data models [12, 16–18, 20]. The data models such as one to one (table mapping), all to one, and selected replication are compared with respect to various NoSQL databases, which are presented in Table 1.

These data models are implemented on various NoSQL databases to measure the performance of each data model.

**Table 1** Various approaches for schema design in NoSQL

S. No.	Data model	NoSQL database		
		Document oriented	Column oriented	Graph based
1.	One to One (table mapping)	Directly maps relation table into collection	Directly maps relational table into column-oriented table	Directly maps tuple of relational table into a node of graph database, and each node has only that key-value pair, which are in a tuple of relational table. Here key means column name of tuple, and makes a relation between nodes by an edge
2.	All to one	Create a single collection, which has different nested documents for each related table or linked table of relation table	Create a column-oriented table, which has different column families for each related table or linked table of relational database	Create a single node, which has all other related node's information in it as key-value pair
3.	Selected replication	Duplicate only those key-value pair as nested document, which are relevant, and required for faster query processing	Duplicate only that information (key-value) of other table, which are relevant, and required for faster query processing in columnar table	Duplicate related key-value pair of one node into other node by examining foreign key in relational tables

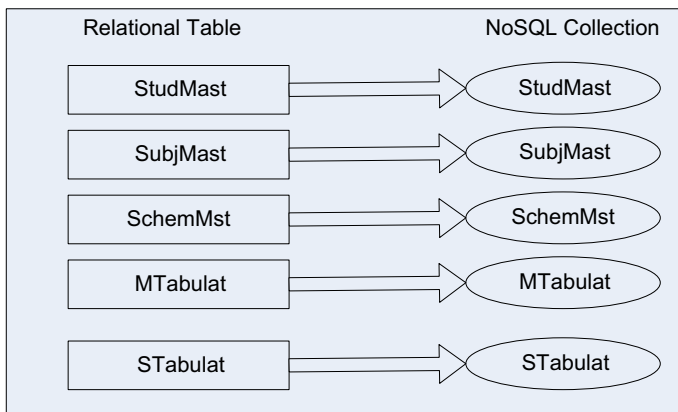
## 4 Experiment Analysis

In order to analyze the performance of above schema design approaches, we have implemented them in various NoSQL databases. We have used MongoDB for document database, Apache Cassandra for column family, and Neo4 J database for a graph database. We have considered our department's result database as a relational database to migrate it in different database schema of NoSQL database. The different query operations are performed in order to find the suitability of schema design in NoSQL database. The select, insert, update, and delete query are executed on each of the schemas. These queries are executed in mongo-shell for MongoDB database, cql shell for Cassandra database, and cipher-shell for Neo4j database. The relational database tables are shown in Table 2.

**Table 2** Relational tables for result database

Table name	Table columns
StudMast	stu_no int, stu_rollno varchar(50), stu_title varchar(10), stu_name varchar(100), stu_sex varchar(1), marital_status varchar(1), nationality varchar(20), fat_hus_name varchar(50), mother_name varchar(50),join_dt datetime,course_cd int,branch_cd int, unv_enroll_no varchar(20),stu_active varchar(1),cur_termno int, result_scheme varchar(5)
STabulat	stab_id int primary key,mtab_id int,sub_srno int, sub_cd varchar(10), int_obt int, int_res text, ext_obt int,ext_res text,trans_score float, sub_grade text,sub_gradep int,score_carry text,repeat_main text
MTabulat	mtab_id int, sess_id int,sch_ctrlno int,stu_no int,stu_stat text,attempt_no int,obt_gradp int,tot_credit int,sem_gpa float,cum_gpa float, stu_pcnt float,stu_division text,gn_entry_flag text,tab_complete text,valid_credit int
SchemMst	course_cd int, branch_cd int, term_no int, sch_year int,sch_ctrlno int, sch_name text, sch_dtl text, sch_active text
SessInfo	sess_id int, primary key, sess_name text, sess_active text, prev_sess int
SubjMast	sub_cd varchar(10),sub_name varchar(100), sub_type varchar(1), sub_credit int, sub_total int, int_total int, ext_total int
ExamMark	sess_id int, sch_ctrlno int,stu_no int, sub_cd varchar(10), int_obt int,ext_obt int

In these tables, we have created three database models, as shown in Fig. 1 for one-to-one table mapping, in which we have created a single document or a table (column family) or a node for its corresponding relational table. Figure 2 explains all-to-one table model, in which all linked related tables of relational database are grouped together in single document (Mongodb) or in single big column family (Cassandra), such as StudMast document contains MTabulat data and MTabulat document contains STabulat, SessInfo, etc. Figure 3 explains the concept of selected replication data model, in which only selected data are ground in single document



**Fig. 1** One-to-one table mapping model

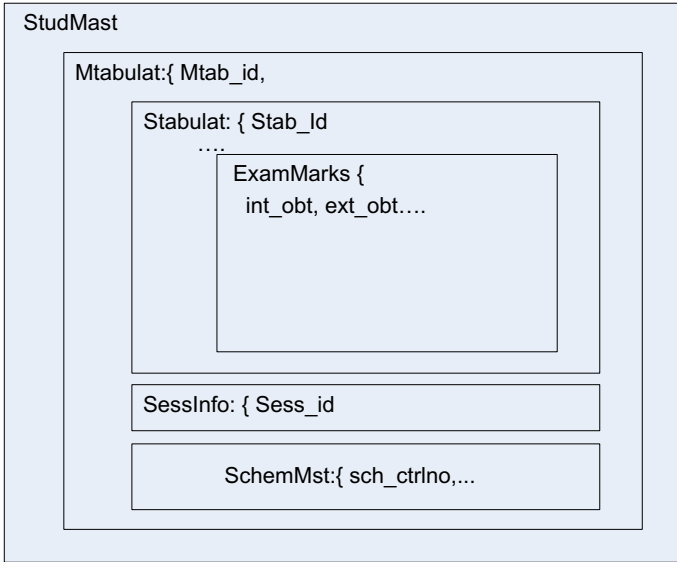


Fig. 2 All-to-one table mapping model

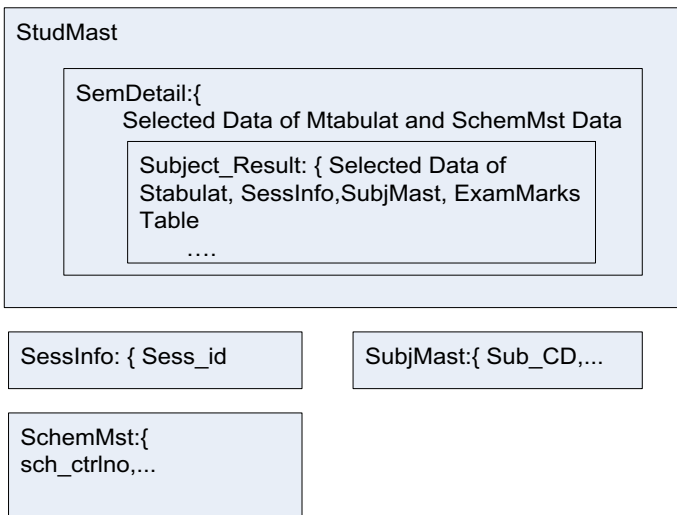
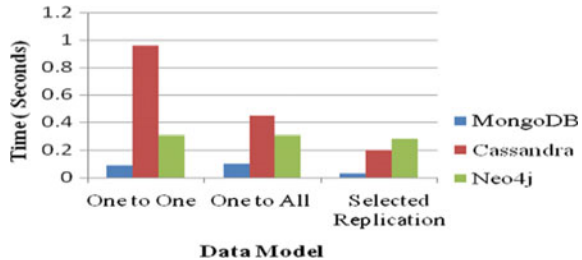


Fig. 3 Selected replication model

**Table 3** Select query information

S. No.	Query details
1.	Find all girls who have SGPA > 8
2.	Find Max of SGPA with Student Details and scheme name in each scheme
3.	Find a particular student record by its roll no
4.	Find all fail student in a particular scheme
5.	Find all students master data who are enrolled in a particular scheme

**Fig. 4** Comparison for select query



(Mongodb) or in single column family (Cassandra) or in single node (Neo4j). In this model, other information may be replicated in separate tables or in separate entities, such as SessInfo, SubjMast and data is also stored in separate table as well as some data is stored in StudMast document under SemDetail sub-document.

We have used select, insert, update, and delete queries, and executed on different NoSQL databases, which have same data values as in relational database. In the subsequent subsections, we are showing the result and performance of them.

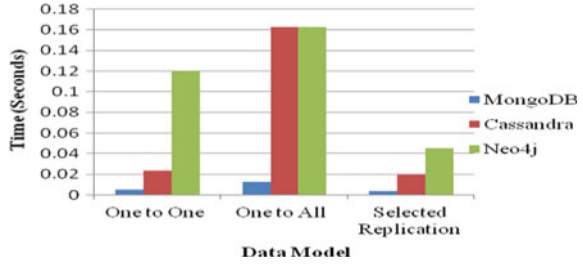
### 4.1 Select Query

The performance of select query is analyzed on different data models of various databases by executing various fundamental select queries on it. These queries are explained in Table 3. The select sample queries for each data model and each NoSQL database-wise query for finding a particular student’s record by its roll number are executed. Figure 4 shows time spent in executing bunch of all five select queries on each system.

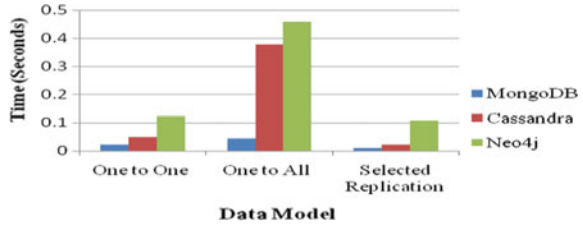
### 4.2 Insert Query

We have analyzed the performance of insert query on different data models of various databases by executing various fundamental insert queries on each model of each

**Fig. 5** Comparison for insert query



**Fig. 6** Comparison for update query



database. We have inserted sample data, such as student master entry in StudMast table and examination marks in ExamMark table, in each data model. Figure 5 shows time spent in executing bunch of insert queries on each system.

### 4.3 Update Query

The performance of update query is analyzed on different data models of various databases by executing various fundamental update queries on each model of each database. We have updated sample data in each data model. Figure 6 shows time spent in executing various update queries in various databases. We also used spark for updating nested data. In spark, we first extract the nested data in some temporary variables and then update it accordingly, and then add/insert it into original Cassandra data table.

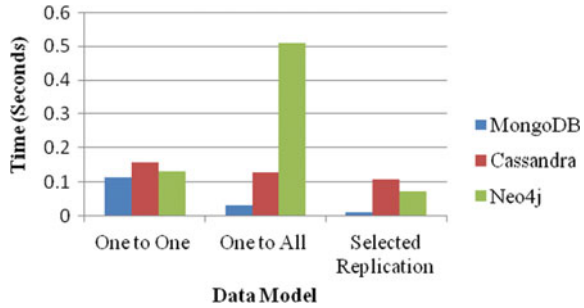
### 4.4 Delete Query

We have analyzed the performance of delete query on different data models of various databases by executing various fundamental delete queries on each model of each database. We deleted sample data in each data model. Figure 7 shows time spent in executing various delete queries in various databases.

On analyzing the performance of select, insert, update, and delete queries on various databases of NoSQL, it is observed that selected replication data model



**Fig. 7** Comparison for delete query



is better than one-to-one and one-to-all data model. Selected replication faces the difficulties for finding the information such as which columns of relational tables have to be duplicated into other NoSQL entity.

## 5 Conclusion

In this paper, performance analysis of various data models, that is, one to one, all to one, and selected replication of NoSQL database, is performed. A departmental university database is considered for the performance analysis. Experimental result shows that NoSQL databases perform better in those schema, where we have denormalized relational tables and grouped related information into single document (document-based database) or single big table (columnar database) or node (graph database). Also, it is observed that the selected replication consumes less time as compared to other two schema models in NoSQL databases, i.e., MongoDB, Cassandra, and Neo4j.

## References

1. Gilbert, S., & Lynch, N. (2002). Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services. *ACM SIGACT News*, 33(2), 51–59.
2. Cattell, R. (2010). Scalable SQL and NoSQL data stores. *ACM Sigmod Record*, 39(4) 12–27.
3. Moniruzzaman, A. B. M. & Hossain, S. A. (2013). NoSQL database: New era of databases for big data analytics—Classification, characteristics and comparison. *International Journal of Database Theory and Application*, 6(4).
4. Sadalage, P. J., & Fowler, M. (2013). *NoSQL distilled: A brief guide to the emerging world of polyglot persistence*. Addison-Wesley Professional.
5. Mior, M. J. (2014). Automated schema design for NoSQL databases. In *Proceedings of 2014 SIGMOD Ph.D. Symposium* (pp. 41–45). ACM.
6. Mior, M. J., Salem, K., Abounaga, A., & Liu, R. (2017). NoSE: Schema design for NoSQL applications. *IEEE Transactions on Knowledge and Data Engineering*, 29(10), 2275–2289.

7. Kuderu, N., & Kumari, V. (2016). Relational database to NoSQL conversion by schema migration and mapping. *International Journal of Computer Engineering in Research Trends*, 3(9), 506–513.
8. Lee, C.-H., & Zheng, Y.-L. (2015). *Automatic SQL-to-NoSQL schema transformation over the MySQL and HBase databases*. Presented at the 2015 IEEE International Conference on Consumer Electronics—Taiwan (pp. 426–427). Taipei, Taiwan, June 6–8, 2015.
9. Chongxin, L. (2010). Transforming relational database into HBase: A case study. In *IEEE International Conference on Software Engineering and Service Sciences* (683–687). Beijing, China.
10. HBase Information. <https://hbase.apache.org/>.
11. Zhao, G., Li, L., Li, Z., & Lin, Q. (2014). Multiple nested schema of HBase for migration from SQL. In *2014 Ninth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC)* (pp. 338–343). Guangdong, China. <https://doi.org/10.1109/3pgcic.2014.127>.
12. Vajk, T., Feher, P., Fekete, K., & Charaf, H. (2013). Denormalizing data into schema-free databases. In *2013 IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom)* (pp. 747–752).
13. Yangui, R., Nabli, A., & Gargouri, F. (2016). Automatic transformation of data warehouse schema to NoSQL data base: Comparative study. *Procedia Computer Science*, 96, 255–264.
14. Stanescu, L., Brezovan, M., & Burdescu, D. D. (2016). *Automatic mapping of MySQL databases to NoSQL MongoDB*. Presented at the 2016 Federated Conference on Computer Science and Information Systems (FedCSIS). Gdansk, Poland, September 11–14, 2016.
15. Karnitis, G., & Arnicans, G. (2015). Migration of relational database to document-oriented database: Structure Denormalization and data transformation. In *7th International Conf. on Computational Intelligence, Communication Systems and Networks (CICSyN)* (pp. 113–118). Riga, Latvia.
16. Zhao, G., Lin, Q., Li, L., & Li, Z. (2014). Schema conversion model of SQL database to NoSQL. In *9th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing* (pp. 355–362). Guangdong, China.
17. Yoo, J., Lee, K.-H., & Jeon, Y.-H. (2018). Migration from RDBMS to NoSQL using column-level denormalization and atomic aggregates. *Journal of Information Science and Engineering*, 34(1), 243–259.
18. <https://neo4j.com/developer/guide-importing-data-and-etl/>.
19. Park, Y., Shankar, M., Park, B.-H., & Ghosh, J. (2014). Graph databases for large-scale health-care systems: A framework for efficient data management and data services. In *IEEE 30th International Conference on Data Engineering Workshops* (pp. 12–19). Chicago, IL, USA.
20. De Virgilio, R., Maccioni, A., & Torlone, R. (2013). Converting relational to graph databases. In *First International Workshop on Graph Data Management Experiences and Systems*. ACM.

# A Time Delay Neural Network Acoustic Modeling for Hindi Speech Recognition



Ankit Kumar and R. K. Aggarwal

**Abstract** Automatic Speech Recognition (ASR) systems have become more popular recently for low resource languages. India has 22 official language and more than two thousands other regional languages, the majority have low resources. The standard resources are also limited for the Hindi language. In this paper, the implementation of continuous Hindi ASR system has been done using Time Delay Neural Network (TDNN) based acoustic modeling significantly improves the performance of baseline Gaussian Mixture Model-Hidden Markov Model (GMM-HMM) based Hindi ASR system up to 11%. Further improvement of 3% and 2% have been recorded by applying i-vector adaptation, interpolated language modeling in this work.

**Keywords** Automatic speech recognition · Acoustic modeling · TDNN · Language modeling

## 1 Introduction

People can Speak their native language very naturally and effortlessly, which can be processed with immense pace and easiness. Speech recognition can be employed in many areas and problems to prove their superiority as an important component of computer interface. Automatic speech recognition is the procedure of parameterization of speech signal at front-end and likelihood estimation at back-end [1]. With the advancement of technology, ASR becomes a more challenging issue because

---

A. Kumar (✉) · R. K. Aggarwal  
Department of Computer Engineering, National Institute of Technology,  
Kurukshetra, Haryana, India  
e-mail: [anketvit@gmail.com](mailto:anketvit@gmail.com)

R. K. Aggarwal  
e-mail: [rka15969@gmail.com](mailto:rka15969@gmail.com)

A. Kumar  
Computer Science and Engineering Department, Galgotias University, Greater Noida, India

of the large-scale real-world applications acoustic environment is much difficult or different than in the past [2].

Acoustic modeling is the key component of any ASR system. The Development of the ASR system for the real-world task was made possible with expectation maximization (EM) algorithm along with the generative method such as GMM-HMM [3]. The Accuracy of GMM-HMM system can further be improved by discriminative training and augmenting the feature (e.g. MFCC) tandem or bottleneck features generated using Neural Network (NN) [4, 5]. Artificial Neural Network (ANN) has the potential to learn and build models for overcoming the inefficiency of GMM [3]. TDNN has been proposed as an alternative of conventional HMM and GMM-HMM [3], subspace GMM [6], Deep Neural Networks (DNN) [7–9], convolutional NN [9–12], RNN [13, 14] and matter of deep investigation. To solve the Hindi ASR problem, we investigate the TDNN based acoustic modeling in this work.

Deep learning and Recurrent Neural Network (RNN) have fueled language modeling research in the past few years. Simpler models, such as n-grams, only use a short history of previous words to predict the next word, they are still a key component to high quality and low perplexity language models [15]. In this work, n-gram language modeling has been used. To train the language model, we have used SRILM [16] toolkit. The interpolation of two different n-gram language model is also done and evaluated on Hindi ASR system.

This paper is organized as follow: Section 2 covers the details of the major part of the ASR system, Sect. 3 describes the Hindi dataset used in this work, Sect. 4 reports the results and Sect. 5 concludes the work.

## 2 System Description

ASR is a practice of converting input speech waveform to word sequence as accurate and efficient as possible. ASR is a very complex problem in which ASR system has to find the most appropriate word sequence  $W^*$  associated with a given acoustic observation sequence  $O$ . This problem can be formulated by the Bayesian framework.

$$W^* = \operatorname{argmax}_w P\left(\frac{W}{O}\right) \quad (1)$$

Using Baye's formula, representation of above equation:

$$P(W|O) = \frac{P(O|W)P(W)}{P(O)} \quad (2)$$

In Eq. 2, the denominator  $P(O)$  s dropped as it is constant and does not change the result of maximization.

$$P(W|O) = P(O|W)P(W) \quad (3)$$

$P(O/W)$  is calculated by the acoustic model and  $P(W)$  is determined by the language model.

## 2.1 Feature Extraction

The process of extracting useful information and discarding others is known as feature extraction [17]. This phase requires more attention as any loss of information can't be retrieved later. Many feature extraction techniques have been developed and used in ASR, out of which MFCC is more common. In this work, we used 39 ( $13 + \Delta + \Delta\Delta$ ) MFCC features to train the acoustic model.

## 2.2 Acoustic Modeling

In this work, we focused on TDNN based acoustic modeling. The acoustic modeling has been done by using available training transcription of Hindi language. The baseline system was trained using GMM-HMM acoustic modeling. GMM-HMM-based acoustic models were based on Maximum Likelihood estimation criterion. The tri-phone based GMM-HMM acoustic modeling was used to generate the alignment for training the neural network models.

### 2.2.1 TDNN Architecture

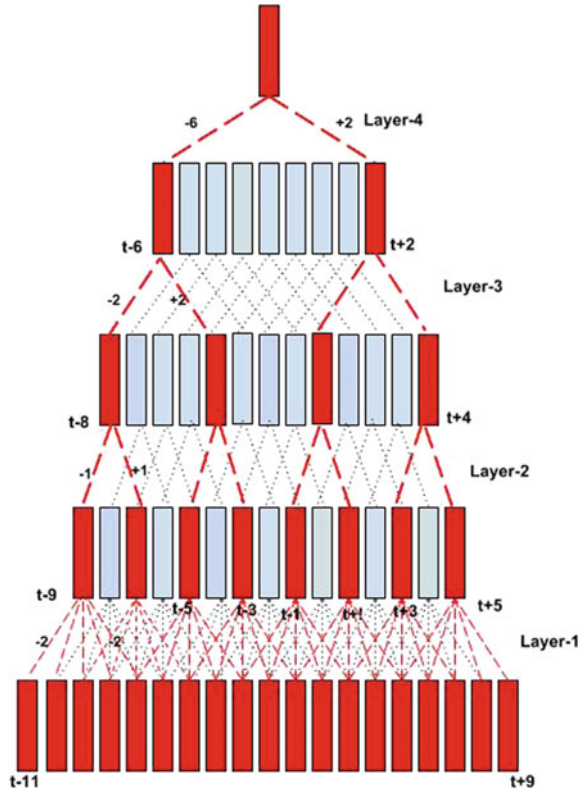
TDNN is a feedforward neural network, which is able to effectively model the long term temporal context [18]. In TDNN architecture, the input layer learns the initial transform on a narrow context and as the hidden layer increases, it performs activation from a wider temporal context [19]. TDNN networks are computationally heavy as hidden activations are computed at every time steps by splicing together contiguous frames in each hidden layer. This computation load can be reduced by applying subsampling, in which splicing is done only for two frames [20].

In Fig. 1, frame  $t-2$  to  $t+2$   $\{-2, -1, 0, +1, +2\}$  or  $[-2, +2]$  are splice together at input layer. In the above mention table, notation  $\{-2, +2\}$  indicate splicing of two frames  $t-2$  and  $t+2$  only to apply activation function. Due to selective activation computational load significantly reduced (Table 1).

## 2.3 Language Modeling

Language model plays a vital role in speech recognition by predicting the next most probable word sequence with the help of preceding  $n-1$  words [21]. In this work, the

**Fig. 1** TDNN-C Architecture based on subsampling



**Table 1** Context specification of TDNN with and without subsampling

Layers	Input context	Input context with subsampling
1	$[-2, +2]$	$[-2, +2]$
2	$[-1, +1]$	$[-1, +1]$
4	$[-2, +2]$	$[-2, +2]$
3	$[-6, +2]$	$[-6, +2]$
3	$\{0\}$	$\{0\}$

text transcription around one million words were collected from the various sources to train the language model. The SRILM toolkit [16] was used to train the language. This work includes the bi-gram, 3-gram, 4-gram, and interpolated 3g + 4g language model to measure the performance of Hindi ASR system.

## 2.4 Pronunciation Lexicon

The dataset used in this work contains the pronunciation of 2803 unique words. Many words in the vocabulary have more than one pronunciation. A total of 68 phones were used to create the pronunciation lexicon. All experiments in this work are based on closed vocabulary, which indicated no use of out of the vocabulary words in test data.

## 2.5 Tools and Performance Measure

The ASR system was implemented using the kaldi toolkit [22]. The Switchboard recipe was used to extract the features, training, and decoding. For language modeling, SRILM toolkit [16] was used. The ASR performance was measured on Word Error Rate (WER).

## 3 Corpus Description

All experiment in this work has been conducted on Hindi dataset by TIFR, Mumbai [23]. The dataset is well annotated and phonetically rich. The dataset contains the speech utterances of 2.5 h duration. The dataset contains 100 speaker utterances. Each speaker utter 10 sentences, out of which 2 sentences remained common. The dataset was recorded in a quiet environment on 16 Khz sampling rate. For training purpose, we randomly select 80 speakers, which include 55 male and 25 female utterances. The remaining were used for testing purpose (Table 2).

## 4 Implementation and Results

All experiments were conducted on Hindi dataset using kaldi toolkit [22]. The training and testing condition remain the same in all experiments. The baseline Hindi ASR system was trained using context-dependent triphone HMM-based acoustic modeling. A total of 68 HMM of Hindi phones was used to train the baseline system. The standard 39 MFCC features were used to train the acoustic model. The bi-gram

**Table 2** Hindi speech corpus details

Dataset	No of speakers	Utterances	Total words	Unique words	Hours
Train	80	800	5420	2015	2.1
Test	20	200	1240	856	0.20

language modeling has been applied by default in all experiments. For language modeling, SRI Language Modeling (SRILM) toolkit has been used. The WER of baseline GMM-HMM system was 33%.

#### 4.1 Experiment with TDNN Acoustic Modeling and I-Vector Adaptation

In this experiment, TDNN based acoustic modeling has been applied on baseline Hindi ASR system. The performance of Hindi ASR has been measured on a different type of TDNN architecture reported in an earlier section. For training, 40 MFCC features were used. In TDNN based acoustic modeling, speed perturbation has been taken place to increase the size of training data. In this work, three fold data augmentation has been applied which create the two more copy of training data with a speed factor of 0.9 and 1.1. The training was done with and without i-vector adaptation. The size of i-vector in this experiment was 100 dimension. In this work, i-vector were estimated in an online fashion during training and during decoding, i-vector were estimated in an off-line fashion.

To train the TDNN, first Universal Background Model (UBM) was created by using all the available data. After the UBM model creation, i-vectors were estimated. At last, TDNN training has been taken place. The initial learning rate was chosen as 0.003000. After several iterations, the final learning rate become 0.00145. We used the chain model recipe of switchboard database in kaldil toolkit. Table 3 clearly shows that TDNN-C architecture performs well in every case. The best WER 16.8% was reported by TDNN-C architecture with i-vector adaptation in 5 epochs.

**Table 3** Results with different TDNN Architectures and i-vector adaptation

SN	Model	Features	i-vector	WER %		
				Epoch 3	Epoch 4	Epoch 5
1	TDNN-A	MFCC-40	No	22	21.5	20.6
2	TDNN-B	MFCC-40	No	21	20.6	20.1
3	TDNN-C	MFCC-40	No	20.5	20.1	19.6
4	TDNN-A	MFCC-40	Yes	19	18.2	17.7
5	TDNN-B	MFCC-40	Yes	18.5	17.9	17.2
6	TDNN-C	MFCC-40	Yes	18	17.4	16.8



**Table 4** Experiment with different language modeling

SN	Acoustic model	WER %			
		bi-gram	tri-gram	4-gram	3g+4g
1	TDNN-A+i-vector	17.7	17.1	16.3	15.9
2	TDNN-B+i-vector	17.7	17.1	16.3	15.9
3	TDNN-C+i-vector	17.7	17.1	16.3	15.9

## 4.2 Experiment with Different Language Modeling

In this experiment, further reduction in WER has been done by applying various language modeling techniques. The SRILM toolkit was used to train the language model. The Kneser-Nay smoothed bi-gram, tri-gram, and 4-gram language models were used to measure the performance gain in this experiment (Table 4).

The interpolation of tri-gram and 4-gram language model was also done to reduce the perplexity. The further 2% relative improvement was reported by TDNN-C with i-vector adaptation using interpolated (3g + 4g) language model. Approx one million text transcription from various sources was used in this experiment.

## 5 Conclusion

Using a combination of TDNN acoustic modeling, i-vector adaptation, and interpolated language modeling, this paper reports a significant reduction in WER. The appropriate section of subsampling indices speeds up the training as well as system performance. The relative 3% improvement has been recorded with i-vector adaptation. The interpolated language model also increases the performance by around 2%. In summary, we found TDNN acoustic modeling is a good choice for Hindi speech recognition. This work can be further explored in different language model and the acoustic model.

## References

1. Aggarwal, R. K., & Dave, M. (2012). Integration of multiple acoustic and language models for improved Hindi speech recognition system. *International Journal of Speech Technology*, 15(2), 165–180.
2. Deng, L., Wang, K., Acero, A., Hon, H. W., Droppo, J., Boulis, C., et al. (2002). Distributed speech processing in MiPad's multimodal user interface. *IEEE Transactions on Speech and Audio Processing*, 10(8), 605–619.
3. Vegesna, V. V. R., Gurugubelli, K., Vydana, H. K., Pulugandla, B., Shrivastava, M., & Vuppala, A. K. (2017). Dnn-hmm acoustic modeling for large vocabulary Telugu speech recognition. In *International Conference on Mining Intelligence and Knowledge Exploration* (pp. 189-197). Springer, Cham.

4. Malioutov, D. M., Sanghavi, S. R., & Willsky, A. S. (2010). Sequential compressed sensing. *IEEE Journal of Selected Topics in Signal Processing*, 4(2), 435–444.
5. Ji, S., Xue, Y., & Carin, L. (2008). Bayesian compressive sensing. *IEEE Transactions on Signal Processing*, 56(6), 2346–2356.
6. Mohan, A., Rose, R., Ghalehjegh, S. H., & Umesh, S. (2014). Acoustic modelling for speech recognition in Indian languages in an agricultural commodities task domain. *Speech Communication*, 56, 167–180.
7. Povey, D., Peddinti, V., Galvez, D., Ghahremani, P., Manohar, V., Na, X., et al. (2016). Purely sequence-trained neural networks for ASR based on lattice-free MMI. In *Interspeech* (pp. 2751–2755).
8. Yoshioka, T., & Gales, M. J. (2015). Environmentally robust ASR front-end for deep neural network acoustic models. *Computer Speech & Language*, 31(1), 65–86.
9. Abraham, B., Umesh, S., & Joy, N. M. (2016). Overcoming data sparsity in acoustic modeling of low-resource language by borrowing data and model parameters from high-resource languages. In *Interspeech* (pp. 3037–3041).
10. Xiong, W., Wu, L., Alleva, F., Droppo, J., Huang, X., & Stolcke, A. (2018). The Microsoft 2017 conversational speech recognition system. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 5934–5938). IEEE.
11. Abdel-Hamid, O., Mohamed, A. R., Jiang, H., Deng, L., Penn, G., & Yu, D. (2014). Convolutional neural networks for speech recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22(10), 1533–1545.
12. Saon, G., Kuo, H. K. J., Rennie, S., & Picheny, M. (2015). The IBM 2015 English conversational telephone speech recognition system. [arXiv:1505.05899](https://arxiv.org/abs/1505.05899).
13. Graves, A., & Jaitly, N. (2014). Towards end-to-end speech recognition with recurrent neural networks. In *International Conference on Machine Learning* (pp. 1764–1772).
14. Mohamed, A. R., Seide, F., Yu, D., Droppo, J., Stoicke, A., Zweig, G., et al. (2015). Deep bi-directional recurrent networks over spectral windows. In *2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)* (pp. 78–83). IEEE.
15. Jozefowich, R., et al. (2016). Exploring the limits of language modeling. [arXiv:1602.02410](https://arxiv.org/abs/1602.02410).
16. Stolcke, A. (2002). SRILM—an extensible language modeling toolkit. In *Seventh International Conference on Spoken Language Processing*.
17. Kuamr, A., Dua, M., & Choudhary, A. (2014). Implementation and performance evaluation of continuous Hindi speech recognition. In *2014 International Conference on Electronics and Communication Systems (ICECS)* (pp. 1–5). IEEE.
18. Peddinti, V., Povey, D., & Khudanpur, S. (2015). A time delay neural network architecture for efficient modeling of long temporal contexts. In *Sixteenth Annual Conference of the International Speech Communication Association*.
19. Peddinti, V., et al. (2015). Reverberation robust acoustic modeling using i-vectors with time delay neural networks. In *Sixteenth Annual Conference of the International Speech Communication Association*.
20. Peddinti, V., et al. (2015). Jhu aspire system: Robust lvcsr with tdnns, ivector adaptation and rnn-lms. In *IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*. IEEE.
21. Kuamr, A., Dua, M., & Choudhary, T. (2014). Continuous Hindi speech recognition using Gaussian mixture HMM. In *IEEE Students' Conference on Electrical* (pp. 1–5). IEEE: Electronics and Computer Science.
22. Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N., et al. (2011). The Kaldi speech recognition toolkit. In *IEEE Workshop on Automatic Speech Recognition and Understanding* (pp. 1–4).
23. Samudravijaya, K., Rao, P. V. S., & Agrawal, S. S. (2002). Hindi speech database. In *International Conference on spoken Language Processing* (pp. 456–464). China: Beijing.

# A Review on Offensive Language Detection



Rahul Pradhan, Ankur Chaturvedi, Aprna Tripathi  
and Dilip Kumar Sharma

**Abstract** Offensive language, hate speech, and bullying behavior is prevalent during textual communication happening online. Users usually misuse the anonymity available online social media, use this as an advantage, and engage in behavior that is not acceptable socially in actual world. Social media platforms, analytics companies, and online communities had shown much interest and involvement in this field to cope up with this problem by stopping its propagation in social media and its usage. In this paper, we will propose the work done by researchers to form effective strategies for tackling this problem of identifying offense, aggression, and hate speech in user's textual posts, comments, microblogs, etc.

**Keywords** Hate speech · N-gram · Offensive language · tf-idf · Machine learning · Twitter · Offensive language detection · Antisocial behavior online

## 1 Introduction

Abusive and offensive language is the prime concern of technical companies nowadays due to exponential growth in number of Internet users around the world and since these people are from different walks of life and different culture. There is a fine line between hate speech and offensive language, and to detect and differentiate among them is a big challenge. In literature, researchers generally classify the text into three classes:

---

R. Pradhan (✉) · A. Chaturvedi · A. Tripathi · D. K. Sharma  
GLA University, Mathura, India  
e-mail: [rahul.pradhan@gla.ac.in](mailto:rahul.pradhan@gla.ac.in)

A. Chaturvedi  
e-mail: [ankur.chaturvedi@gla.ac.in](mailto:ankur.chaturvedi@gla.ac.in)

A. Tripathi  
e-mail: [aprna.tripathi@gla.ac.in](mailto:aprna.tripathi@gla.ac.in)

D. K. Sharma  
e-mail: [dilip.sharma@gla.ac.in](mailto:dilip.sharma@gla.ac.in)

- Hateful,
- Offensive, and
- Clean

In this paper, we showcase the study we perform on the research held in this area with some light on what can be done next in order to make it more efficient. Our objective behind carrying this work is to come up with a study of papers and research work done in this field so far.

## 2 Terminology

In this paper, we use the term hateful, offensive, and clean. We come to a conclusion in favor of the usage of these terms since they can have broader meaning and can be used in various contexts in user-generated content to define it first. Hateful text or speech is not a very common phrase to refer to such text in legal world but in general terms day-to-day speaking we use it quite often.

Following is the list of terms used in literature [1]:

- abusive messages,
- hostile messages, or
- flames.

This will help readers to go further in literature on this topic. There is a recent trend in the NLP world that author prefers to use the word cyberbullying [2–7].

Hateful speech or hate speech is commonly referred to as conversation, communication that mocks a group of people or a single person on the grounds of social status, race, color, ethnicity, nationality, gender, sexual preferences, religions, and many others [8].

## 3 Literature Survey

Researchers in past have proposed various machine learning approaches and their variant to deal with the problem of offensive language. Detecting sarcasm had been the point of research for many researchers around in area of NLP or text mining, with need of hour nowadays people are more focusing on detecting the wrongs prevailing in social media. This concern of government and public leads to open new research domains as fake news detection, rumor detection, offensive language detection, etc. Many of these proposed works use feature extraction from text such as bag of words (BOW) and dictionaries. Major work in this area is focused on feature extraction from text. Dictionaries [9] and bag of words [10] were among the lexical features that were used widely by researchers to detect the offensive language or phrases.

Gaydhani et al. [11] used tf-idf and N-gram as features for their classification of tweets with 95.6% accuracy.

It was found out that these features could not understand the context of sentences. Approaches that involve N-gram show better results and perform better than their counterparts [12]. Lexical features are proving to outperform other features in automatic detection of offensive language and phrases, without taking into consideration the syntactic structures as bag of word approach could not detect offensiveness if words are used in different sequences [13].

Gaydhani et al. [11] form a dataset which is the combination of three different datasets. The first dataset which they used is publicly available on Crowdflower1, which was used in [14, 15]. Dataset Crowdflower1 has tweets classified into three classes: “Hateful”, “Offensive”, and “Clean”. All the tweets in this dataset are manually annotated. The second dataset they used is crowdflower2 having tweets manually classified into same three classes. Github3 is the third dataset they integrate with other two to build their dataset for study. This third dataset consists of two columns: tweet-ID and class. “Sexism”, “Racism”, and “Neither” are the three categories or classes in which each of these tweets are classified. This dataset is used by [14, 16]. They have considered logistic regression, naive Bayes, and support vector machines for text classification. They used training of dataset on each model by performing grid search for all the combinations of feature parameters and performed 10-fold cross-validation. They analyzed performance on the basis of average score of the cross-validation.

Davidson et al. [17] reduce the dimensionality of the data using a logistic regression with L1 regularization. They show a comparative study on prior work such as logistic regression, naive Bayes, decision trees, random forests, and linear SVMs. They use fivefold cross-validation, with keeping 10% of the sample for evaluation to help prevent overfitting on all the models. Their study suggests that logistic regression and linear SVM perform slightly better than other models. They further use logistic regression with L2 regularization for the final model as it has shown better result in previous work. They use tweets from Hatebase.org which contains lexicon compiled by Internet users containing words and phrases that are considered to be hate speech. Using these words from lexicon they crawled the twitter using the Twitter API which collects tweets containing these words. They collect 33,458 user’s tweets as sample. They get these tweets annotated by CrowdFlower workers into three categories: hate speech, offensive but not hate speech, or neither offensive nor hate speech. Getting these manually annotated helps in clear tagging as they not just look for words but also context of tweets. They found majority of the tweets fall into category of offensive language. They use features from these tweets and used them to train a classifier.

Lee et al. [18] use the dataset titled “Hate and Abusive Speech on Twitter” [19] recently released. This dataset contains the tweets classified into four categories, namely, “normal”, “spam”, “hateful”, and “abusive”.

70 character dimensions using 26 lower character dimensions were used to convert the tweets into one hot encoded vector with 10 digits and special characters up to 34 including whitescape. This encoding is used for character-level representation.

**Table 1** Distribution of categories among tweets

Categories	Normal	Spam	Hateful	Abusive
Number	42,932	9,757	3,100	15,115
(%)	(60.5)	(13.8)	(4.4)	(21.3)

Before this encoding, they have removed user ID, emojis, and URLs, and replace them by special tokens.

Table 1 shows the distribution of tweets among four categories, which are discussed below:

Aken et al. [20] consider two datasets to evaluate their proposed algorithm: one of the datasets they pick from Kaggle’s second challenge on toxic comment classification which contains comments on Wikipedia talk pages presented by Google Jigsaw and other datasets they consider are based on Twitter by Davidson et al. [21]. Class distribution of both datasets is shown in Tables 2 and 3. 24,783 tweets were extracted from Twitter which constitute to dataset of Davidson et al. [21], and all these tweets were annotated by CrowdFlower workers with the labels “hate speech”, “offensive but not hate speech”, and “neither offensive nor hate speech”.

They propose an ensemble to figure out that a single classifier is most effective on certain kind of comment. The ensemble classifier analyzes the features from comments, weights, and for a given feature combination it identifies the suitable single classifier. To attain the goal of identifying the classifier using gradient boosting decision tree, they perform validation across the average final predictions on five trained models.

The most valuable contribution by Aken et al. [20] is Error Classes of False Negatives they have defined. These classes are as such Doubtful labels, and these are the labels that cannot be clearly identified as toxic because for a particular user it is toxic but there are users or annotators that consider it as nontoxic. Second class of false-negative error is Tweets that contain toxicity without any kind of hate words or swear words that this class of error needs to overcome which will require investigating some semantic embeddings for obtaining better classification on different paradigmatic contexts. Third class of error identified by the author is Rhetorical

**Table 2** Wikipedia comment dataset

Categories	Clean	Toxic	Obscene	Insult	Identity hate	Severe toxic	Threat
Number	2,01,081	21,384	12,140	11,304	2,117	1,926	689
%	80.23	8.53	4.84	4.51	0.84	0.77	0.27

**Table 3** Twitter dataset

Categories	Offensive	Clean	Hate
Number	19,190	4,163	1,430
%	77	17	6

Questions and these are the kind of text sentences that does not contain any toxic words but have sarcastic questions in it, usually such text contains question marks and question words. Other classes they introduced are Metaphors and comparisons, and idiom that can be twisted in meaning by looking at context which are difficult to see in short text and such text usually requires knowledge about the implications of language or some additional contextual knowledge. Aken et al. [20] find that different approaches fail in identifying different texts and make errors, but this can be combined into an ensemble with F1-measure. They find some combination of shallow learners with deep neural networks showing remarkable results and proved it to be very effective.

Mathur et al. [22] explore the usage of mixed language in their work and identify the offensive text or hate speech. They choose Hinglish as their subject because of its ease in communication and being popular on Twitter due to its reachability to larger audience in native language. They faced difficulty as this mix of two languages has inherent variations of spellings and absence of grammar induces considerable amount of ambiguity to text and makes the problem even harder to disambiguate and understand the true meaning of text. They proposed the multi-input multichannel transfer learning (MIMCT)-based model is used to identify and detect the hate speeches and offensive language in Hinglish tweets. They use the dataset proposed by them and named it as Hinglish Offensive Tweet (HOT) dataset. Their proposed learning model uses multiple feature inputs using transfer learning. They employed word embedding with secondary extracted features as input to train their multichannel CNN LSTM which is pretrained on English tweets.

Table 4 shows the distribution of tweets among different classes in HOT dataset.

Pitsilis et al. [23] address the effectiveness of identifying the class (being offensive or not offensive) of new tweet or post, using the identity and history of user who has posted the tweet and other tweets posted by him or by other user related to him. They use LSTM for classification and classify the tweets into three classes, namely, neutral, racism, and sexism. The dataset they used is proposed by Waseem et al. [24] and contains about 16,000 short messages collected across Twitter (Table 5).

The biggest issue with this dataset is of dual labeled tweets in the dataset. The number of these tweets is not that small that they can ignore them. Being more precisely, there are 42 tweets that are annotated as both “Neutral” and “Sexism”,

**Table 4** Hinglish offensive tweet (HOT) dataset

Categories	Non-offensive	Abusive	Hate inducing
Number	1121	1765	303
%	35.15	55.35	9.5

**Table 5** Waseem Twitter dataset

Categories	Racism	Sexism	Neutral
Number	1943	3166	10,889
%	12.15	19.79	68.06

while 06 tweets were classified as “Racism” and “Neutral” both. According to the dataset providers, the labeling was performed manually.

Wiedemann et al. [25] explore different techniques for automatic detection of offensive text or hate speech on Tweets written in German language. They also employ deep learning for this task and use a series BiLSTM and CNN neural network in sequence. They improve the accuracy of three learning transfer task for improving the classification performance using context and historical data. They compare supervised categories such as near offensive to weakly supervised categories that contain emojis, and they also show comparison to unsupervised category using tweets of same topic by clustering them using latent Dirichlet allocation (LDA).

## 4 Conclusion

In this paper, we try to present the work done recently in this field of automatic detection of offensive language. We show that how research goes from using tf-idf to popular classifiers such as naïve Bayes, support vector machine (SVM), logistic regression, and then research work goes to variant of these classifiers such as linear SVM, logistic regression with L2, and from here researchers further explore ensemble classifiers using the combination of these classifiers by decomposing the task into subtasks, and then lastly the usage of deep learning and we found many researchers using approaches such as LSTM, CNN, and RNN. Each of these techniques has their own advantages and for classification accuracy LSTM models have outperformed others.

## References

1. Spertus, E. (1997). Smokey: Automatic recognition of hostile messages. In *Proceedings of the Fourteenth National Conference on Artificial Intelligence and Ninth Conference on Innovative Applications of Artificial Intelligence, AAAI'97/IAAI'97* (pp. 1058–1065). Providence, RI, USA: AAAI Press.
2. Xu, J.-M., Jun, K.-S., Zhu, X., & Bellmore, A. (2012). Learning from bullying traces in social media. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 656–666). Montreal, Canada: Association for Computational Linguistics.
3. Hosseinmardi, H., Mattson, S. A., Rafiq, R. I., Han, R., Lv, Q., & Mishra, S. (2015). Detection of cyberbullying incidents on the instagram social network. [abs/1503.03909](https://arxiv.org/abs/1503.03909).
4. Zhong, H., Li, H., Squicciarini, A. C., Rajtmajer, S. M., Griffin, C., Miller, D. J., & Caragea, C. (2016). Content-driven detection of cyberbullying on the instagram social network. In *IJCAI* (pp. 3952–3958). New York City, NY, USA: IJCAI/AAAI Press.
5. Hee, C. V., Lefever, E., Verhoeven, B., Mennes, J., Desmet, B., De Pauw, G., Daelemans, W., & Hoste, V. (2015). Detection and fine-grained classification of cyberbullying events. In *Proceedings of Recent Advances in Natural Language Processing, Proceedings*, Hissar, Bulgaria (pp. 672–680).



6. Dadvar, M., Trieschnigg, D., Ordelman, R., & de Jong, F. (2013). Improving cyberbullying detection with user context. In *Proceedings of the European Conference in Information Retrieval (ECIR)*, Moscow, Russia (pp. 693–696).
7. Dinakar, K., Jones, B., Havasi, C., Lieberman, H., & Picard, R. (2012). Common sense reasoning for detection, prevention, and mitigation of cyberbullying. *ACM Transactions on Interactive Intelligent Systems*, 2(3), 18:1–18:30.
8. Nockleby, J. T. (2000). Hate speech. In L. W. Levy, K. L. Karst, & D. J. Mahoney (Eds.), *Encyclopedia of the American constitution* (pp. 1277–1279, 2nd ed.). Macmillan.
9. Liu, S., & Forss, T. (2015). New classification models for detecting Hate and Violence web content. In *2015 7th international joint conference on knowledge discovery, knowledge engineering and knowledge management (IC3K)*, Lisbon (pp. 487–495).
10. Burnap, P., & Williams, M. (2016). Us and them: Identifying cyber hate on Twitter across multiple protected characteristics. *EPJ Data Science*, 5(1).
11. Gaydhani, A., Doma, V., Kendre, S., & Bhagwat, L. (2018). Detecting hate speech and offensive language on twitter using machine learning: An N-gram and TFIDF based approach. [arXiv:1809.08651](https://arxiv.org/abs/1809.08651).
12. Nobata, C., Tetreault, J., Thomas, A., Mehdad, Y., & Chang, Y. (2016). Abusive language detection in online user content. In *Proceedings of the 25th International Conference on World Wide Web-WWW'16*.
13. Chen, Y., Zhou, Y., Zhu, S., & Xu, H. (2012). Detecting offensive language in social media to protect adolescent online safety. In *2012 international conference on privacy, security, risk and trust and 2012 international conference on social computing*.
14. Watanabe, H., Bouazizi, M., & Ohtsuki, T. (2018). Hate speech on twitter: A pragmatic approach to collect hateful and offensive expressions and perform hate speech detection. *IEEE Access*, 6, 13825–13835.
15. Davidson, T., Warmlesley, D., Macy, M., & Weber, I. (2017). Automated hate speech detection and the problem of offensive language. In *International AAAI conference on web and social media*.
16. Waseem, Z., & Hovy, D. (2016). Hateful symbols or hateful people? Predictive features for hate speech detection on twitter. In *Proceedings of the NAACL Student Research Workshop*.
17. Davidson, T., Warmlesley, D., Macy, M., & Weber, I. (2017). Automated hate speech detection and the problem of offensive language. In *ICWSM 2017*.
18. Lee, Y., Yoon, S., & Jung, K. (2018). Comparative studies of detecting abusive language on twitter. [arXiv:1808.10245](https://arxiv.org/abs/1808.10245).
19. Founta, A., Djouvas, C., Chatzakou, D., Leontiadis, I., Blackburn, J., Stringhini, G., Vakali, A., Sirivianos, M., & Kourtellis, N. (2018). Large scale crowdsourcing and characterization of twitter abusive behavior. In *Proceedings of the International AAAI Conference on Web and Social Media*.
20. van Aken, B., Risch, J., Krestel, R., & Löser, A. (2018). Challenges for toxic comment classification: An in-depth error analysis. [arXiv:1809.07572](https://arxiv.org/abs/1809.07572).
21. Davidson, T., Warmlesley, D., Macy, M. W., & Weber, I. (2017). Automated hate speech detection and the problem of offensive language. In *ICWSM 2017*.
22. Mathur, P., Sawhney, R., Ayyar, M., & Shah, R. (2018). Did you offend me? Classification of offensive tweets in hinglish language. In *Proceedings of the 2nd Workshop on Abusive Language Online (ALW2)* (pp. 138–148).
23. Pitsilis, G.K., Ramampiaro, H., & Langseth, H. (2018). Detecting offensive language in tweets using deep learning. [arXiv:1801.04433](https://arxiv.org/abs/1801.04433).
24. Waseem, Z., & Hovy, D. (2016). Hateful symbols or hateful people? Predictive features for hate speech detection on twitter. In *Proceedings of the NAACL Student Research Workshop, San Diego, California, June 2016*. Association for Computational Linguistics.
25. Wiedemann, G., Ruppert, E., Jindal, R., & Biemann, C. (2018). Transfer learning from LDA to BiLSTM-CNN for offensive language detection in twitter. [arXiv:1811.02906](https://arxiv.org/abs/1811.02906).

# Attribute-Based Elliptic Curve Encryption for Security in Sensor Cloud



Munish Saran, Rajendra Kumar Dwivedi and Rakesh Kumar

**Abstract** Security is one of the major challenges in the field of sensor cloud. There is a need to implement security over sensor cloud using a technique which involves fine-grained access in virtualized wireless sensor networks. The existing security model to secure the data transmission and stored data at the sensor-cloud environment uses different encryption techniques, but its effectiveness, efficiency, and performance can be further increased. Earlier approaches provided fine-grained access such as Ciphertext-Policy Attribute-Based Encryption (CP-ABE) which involves some complex computations. In this paper, we proposed a security model based on Elliptic Curve Encryption (ECC) and attributes to ensure the overall security of sensor data that guarantees the confidentiality and integrity. It also provides a fine-grained access control. The proposed approach reduces the overall computational overhead as compared to other existing approaches.

**Keywords** Data security · Cloud computing · Wireless sensor networks · Sensor cloud · Virtualization · Elliptic curve encryption · Ciphertext-Policy Attribute-Based Encryption · Proxy re-encryption · ElGamal encryption

## 1 Introduction

Sensor cloud is the integration of sensor networks with cloud computing. Security is one among the major several challenges in sensor cloud. Despite several advancement and research toward the security of sensor cloud, there is still a need to focus on this section. Data collected by the sensors are of great value for the end users as the end user's requirement relies heavily on the integrity and accuracy of the data.

---

M. Saran (✉) · R. K. Dwivedi (✉) · R. Kumar (✉)  
Department of CSE, MMMUT Gorakhpur, Gorakhpur, India  
e-mail: [munishsaran@gmail.com](mailto:munishsaran@gmail.com)

R. K. Dwivedi  
e-mail: [rajendra.gkp@gmail.com](mailto:rajendra.gkp@gmail.com)

R. Kumar  
e-mail: [rkiitr@gmail.com](mailto:rkiitr@gmail.com)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_42](https://doi.org/10.1007/978-981-15-0694-9_42)

Data modified by the intruder or hacker will not result in the correct analysis upon data. And also huge loss to Cloud Service Provider (CSP) if detected the forgery of data, the CSP may have to pay penalty as well as loss of trust and reputation between CSP and CSU (cloud service user). Example in the case of medical data collected by various body sensors is very crucial, as any alteration or loss in the medical data of a patient results negatively in the health condition sometimes leading to very serious stages even death of patient. Various security mechanisms are available to provide data security at the cloud. On integrating the sensor network with cloud computing, the security risks further get increased. Hence, there is a need to provide an improved security mechanism for fine-grained access control with reduced complexity. The proposed approach in this paper makes use of attributes to provide fine-grained access control and elliptic curve encryption technique for the purpose of providing enhanced security with reduced complexities.

Rest of the paper is organized as follows. Background details of wireless sensor network and sensor cloud are discussed in Sect. 2. Section 3 describes the related work over security of sensor cloud. Section 4 presents the proposed model for security over sensor cloud along with the algorithm and flowchart. Section 5 presents the performance evaluation and the results and Sect. 6 concludes the paper with some future directions.

## 2 Preliminaries

This section describes the background details for wireless sensor network and sensor cloud and virtualization giving a brief idea about both of them.

### 2.1 *Wireless Sensor Networks*

WSN consists of sensor nodes which are from different vendors and are heterogeneous in nature including sensors of different kinds such as sound sensor, camera sensor, temperature sensor, proximity sensor, motion sensor, light sensor, color sensor, accelerometer sensor, etc. In today's world, the use of WSN has immensely grown helping in various applications [1–5] such as developing smart cities, health-care monitoring, defence and military, security purposes, smart home monitoring, etc.

With ever increase in the demand for WSN applications one cannot increase the number of deployed sensors in the overall system all the time to meet the requirement. Here comes the concept of virtualization of sensors [6]. Virtualization is of great importance to overcome the limitations of WSN providing several benefits such as performance improvement, increase scalability, easy management and maintenance of sensor nodes, reduce cost benefits, and better utilization of resources. Among all the abovementioned benefits increasing the scalability is the one which makes the

overall system more productive as well as profitable. With the exponential increase in the count of IoT applications day by day and their huge dependency on the deployed sensors for data collection, the virtualization of sensors is the only better solution.

## 2.2 *Sensor Cloud*

With various limitations in WSN such as storage, battery power, computation power, data security issues, etc. [7], there is a need to integrate WSN with cloud. As cloud computing provides various powerful features which overcome the shortcomings of WSN by providing huge storage and computational capability, data security, real time data processing, scalability, multitenancy, etc. In simple words, the integration on sensor cloud [8–11] works in a way such that the sensor nodes collect the data and sends to cloud for further processing, allowing access of data from anywhere and at any time instantly [12].

Sensor cloud can be defined as the integration of WSN with cloud computing. Some of the existing sensor-cloud applications include Nimbits, Pachube Platform, IDigi, ThingSpeak.

## 3 Literature Review

Alamri et al. [13] architecture of sensor cloud depicting the issues and challenges for the integration of WSN with cloud computing.

Islam et al. [14] gave the applications of Virtual Sensor Networks (VSNs) such as monitoring the battleground, smart house monitoring and monitoring animal crossing, structural monitoring, health care, agricultural monitoring, and industrial monitoring.

Thilakanathan et al. [15] suggested a platform for the sake of secure monitoring as well as sharing of health data in cloud using ElGamal-based proxy re-encryption method for implementing the security protocol.

Tu et al. [16] proposed revocation mechanism which provides access control in fine-grained manner. This paper addresses the major problems of sharing the data in cloud as well as removing the access rights from the same user when he/she is not the part of the system concerned using CP-ABE [17].

Li et al. [18] proposed a security mechanism for the purpose of securing the health records while sharing by using ABE. The framework described in this consists of predefined list of authorized users who are allowed to access the health records of the patients including medical professionals as well as family members [19]. Attributes based on roles are assigned to each user and the corresponding secret key is also retrieved from the authority and distributed to the user. Role-based policy provides better key management facility for the users. Also this framework is much more

effective than the previously proposed approach as it allows the data owners need not be online all the times.

Tran et al. [20] gave a framework which states that same group users can access the data of each other. This helps in data sharing among the group members. There is a group administrator who is responsible for the revocation of group members. This framework uses proxy re-encryption in which the private key of the data owner is divided into two halves. The first part is stored on the proxy through which the complete data is encrypted. The second part of the private key is kept in the machine of data owner by which he encrypts the data.

Hung et al. [21] gave multiuser data encryption scheme through multiple proxies instead of just single proxy. Separate storage as well as query keys is provided to every user. This makes the queries of the user to remain unrevealed to other user and attacks. And both these storage and query keys can be changed by the user even without decrypting the complete database.

Bethencourt et al. [22] gave a scheme known as ciphertext-policy attribute-based encryption. In this encryption technique, access policies are defined and if the attributes satisfy these access control policies the data will be made available to the user. There are two types of this encryption—KPABE (based on key policy) and CP-ABE (based on ciphertext policy). In CP-ABE, the control policy is associated with data while the policy attributes with the key. KP-ABE is vice versa to CP-ABE in its working. The user's private key stores the access policy while the attributes with data.

Sayantani et al. [23] gave the model for security in the sensor-cloud integration environment which helps in accessing as well as managing the data collected by WSN [24–26] taking security, communication, and processing into account. Security can be made efficient by taking the care for core aspects such as availability, integrity (through SHA algorithms), confidentiality, and access control (through cryptography technique like AES) (Table 1).

## 4 Proposed Model for Security Over Sensor Cloud

The proposed model for security over sensor cloud can be used within the scenarios of virtualized WSN environment. There are several cases of virtualized WSN such as health care, smart house monitoring, industrial monitoring, battlefield monitoring, structural monitoring, rock slides, animal crossing monitoring, and agricultural monitoring. Efficient security mechanism in these scenarios is very crucial as only the authorized users must be allowed to access the data and this is due to the presence of many different types of actors/users in any system.

This architecture consists of the following four entities: data owner (sensor data collector via WSN), security authority (which specifies the security policies), data consumers (end users), and storage servers (cloud). The security authority provides all the attributes for the entire organization to the data owner under which he/she can make the policy by clearly stating who can access what. This policy is made as the

**Table 1** Summary of related work

Author	Proposed approach	Pros	Cons
Thilakanathan et al. [15]	ElGamal encryption	Stable in case of large data sizes and user revocation	Based on the assumption that the data sharing party is fully trusted
Tu et al. [16]	Based on dual encryption system	Very efficient in revoking the access rights from the users	Not suitable for very large data size sets
Li et al. [18]	Attribute-based encryption (ABE)	Supports dynamic modification of file attributes and access policies	Suffers from complex computational overheads in the ABE
Tran et al. [20]	Proxy re-encryption	Allows users to gain access over another user data who are in the same group	Proxy suffers from many encryption and decryption operations
Nguyen et al. [21]	Proxy re-encryption technique employing ElGamal as well as bilinear map	The framework allows multiple users to access the shared database securely	More number of proxies are used
Yang et al. [27]	Both proxy re-encryption and attribute-based encryption	Efficient in the case of simple user revocation scheme	Fails in the scenario when a revoked user joins the group again with a different access privilege

combination of attributes over conjunction or disjunction. After making the policy, the data owner makes use of encryption key provided by the security authority and encrypts the file containing data. This encrypted file is attached along with the policy and then uploaded to the cloud. The attributes of data consumer are registered with the security authority at the time of registration of data consumer. Data consumer selects the file for its usage but the file will be visible only if his/her attributes satisfy the access policy created over that particular file. If satisfied the data consumers can decrypt the file using the decryption key provided by the security authority. Figure 1 describes the proposed architecture and its flowchart is given in Fig. 2.

There are four essential modules involved in this proposed approach, namely, security authority, data owner, data consumer, and cloud database. The security authority works as the admin for the application. He/she is solely responsible for the creation of necessary attributes which are required for the entire system under observation. The role for data owner is to collect the sensor data through various sensors and also create role-based policy based on attributes provided by the security authority for the authentic data access. After this step, the data is encrypted through key provided by security authority and stored on the Azure cloud database. Only the authentic data consumers whose attributes satisfy the security policy for the selected file will be able to decrypt the file with the decryption key provided from the security authority (Table 2).

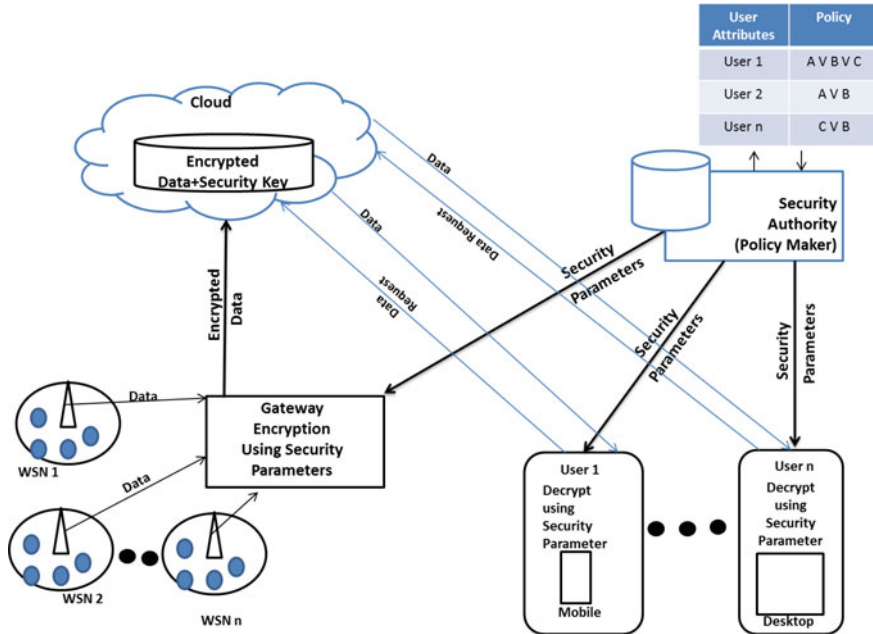


Fig. 1 Proposed model for security over sensor cloud

## 5 Performance Evaluation

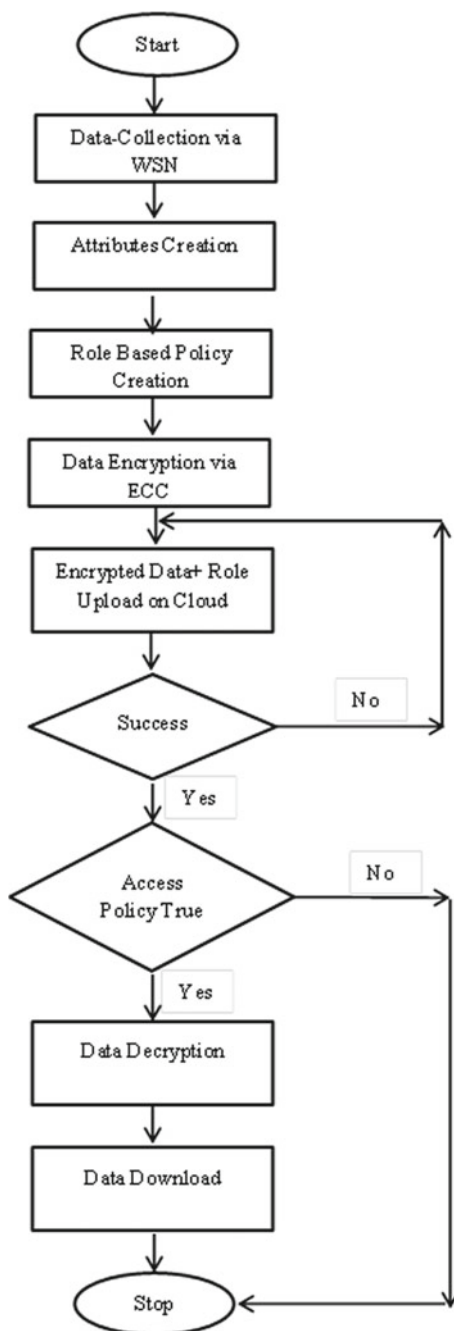
### 5.1 Implementation Setup

This section highlights the implementation of our proposed approach. Finally, performance testing is performed in order to test the overall performance as well as stability of our solution. Table 3 describes the implementation details.

### 5.2 Encryption Time Analysis

Fine-grained access control mechanism [28–30] is provided by CP-ABE [31]. But this CP-ABE policy has complex processing overheads in terms of encryption, decryption, and key-generation time. In order to overcome these complexities, our proposed solution makes use of Elliptic Curve Cryptography (ECC) [32–34] for the encryption of the data and at the same time attaching the encrypted data with role/attributes provided by the security authority. The encryption analysis is shown in Fig. 3. The uploading files used were of varying sizes 10, 20, 30, 40, and 50 KB.

**Fig. 2** Flowchart for attributes-based ECC





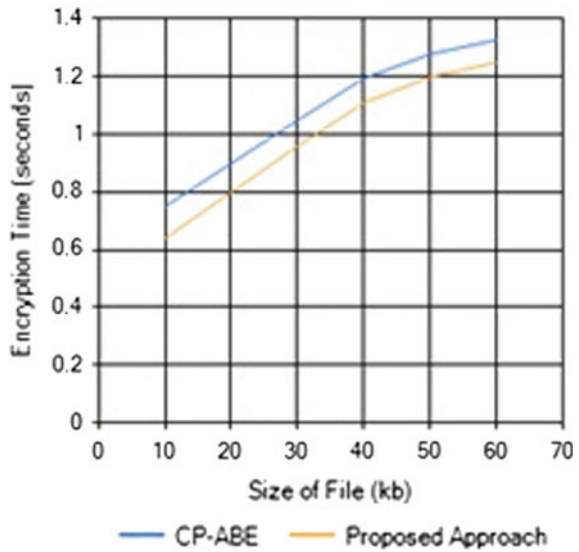
**Table 2** Algorithm for attribute-based elliptic curve encryption

Algorithm: attribute-based elliptic curve encryption	
1	Begin
2	Security Authority registers into the system (as System Admin)
3	System Admin login the system with valid credentials
	System Admin describes the list of attributes for the concerned organization
4	Data Owner registers into the system
5	Data Owner logins using his/her credentials
6	Data Owner defines the access policy from the list of attributes defined by the security authority
	WSN collects sensor data
7	Data owner choose the file that containing the data collected by WSN
8	Data owner encrypts the selected file using the encryption key provided by the security authority
9	Data gets encrypted using Elliptic Curve Cryptography (ECC)
10	Data owner uploads the encrypted file along with the selected policy (attributes) on the cloud server
11	Data owner gets the confirmation email (success/failure) as the file gets uploaded on the cloud server
12	Data Consumer registers into the system
13	Data Consumer logins using his/her credentials
14	Data Consumer selects the file for download
15	Data Consumer will be able to access the file if the access policy over that data file defined by the data owner is satisfied by his/her attributes
16	Data Consumer decrypts the chosen file using the decryption key provided by the security authority
17	File gets downloaded on the data consumer's application end
18	End

**Table 3** Implementation environment

Implementation environment	
Application type	Web application
IDE used	Visual Studio 2012
Framework	Microsoft .NET Framework Version 4.0
Technology	C#, ASP.NET, ADO.NET, JQuery
Backend database	Microsoft SQL Server Management Studio 2012
Cloud database	Microsoft Azure SQL Database
Modules involved	Security Authority, Data Owner, Data Consumer, Cloud Database

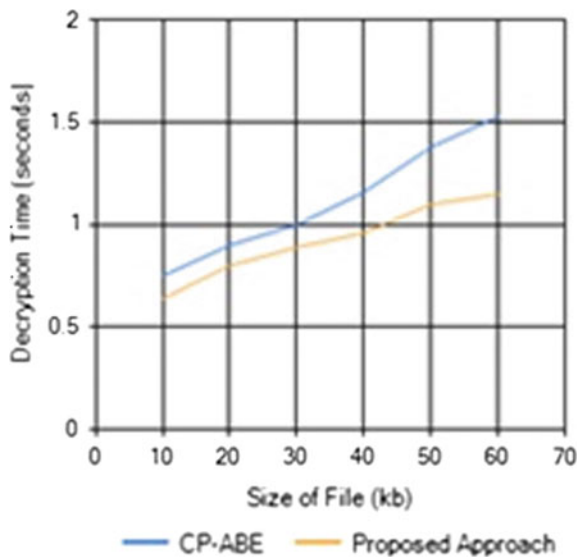
**Fig. 3** Encryption time analysis



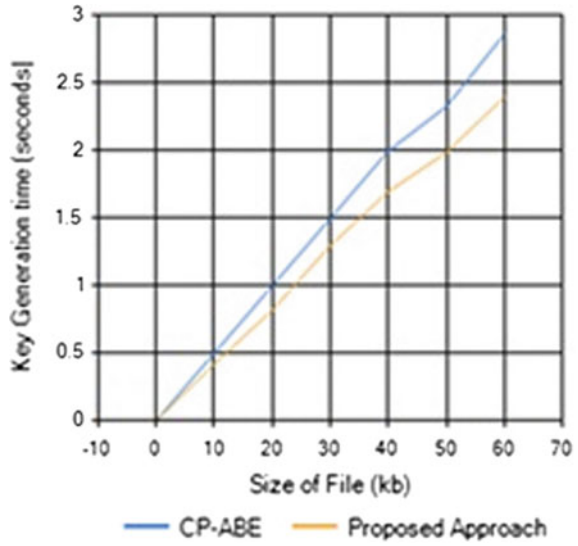
### 5.3 Decryption Time Analysis

The varying size files of size 10, 20, 30, 40, and 50 KB were used for the purpose of evaluating the time needed for decrypting those files by existing CP-ABE [35] scheme as compared with the proposed scheme. It is observed that the decryption time is less in our scheme than that needed in the traditional scheme. Figure 4 clearly shows the analysis.

**Fig. 4** Decryption time analysis



**Fig. 5** Key-generation time analysis



**Table 4** Summary of complexity analysis

Size of file (Kb)	Encryption time (s)	Decryption time (s)	Key-generation time (s)
10	0.64	0.5	0.5
20	0.8	0.6	1
30	0.96	0.7	1.5
40	1.11	0.8	2.0
50	1.2	0.9	2.33
60	1.25	1.0	2.866

### 5.4 Key-Generation Time Analysis

The varying size files of size 10, 20, 30, 40, and 50 KB were used for the purpose of evaluating the time needed for key generation for those files by existing CP-ABE scheme as compared with the proposed scheme. It is observed that the key-generation time is less in our scheme than that needed in the traditional scheme and this is referred in Fig. 5 (Table 4).

## 6 Conclusions and Future Directions

Sensor data are very critical and its security cannot be compromised at any cost as well as at any stage. Various security policies and algorithms can be applied over the sensor data to protect it. Due to the virtualization of WSN, the need for securing the data even further increases, as the data collected from the virtualized environment must be safely delivered to the intended users. This is achieved in our proposed architecture by making use of ECC which encrypts the sensor data on the basis of attributes by the data owner and only those users who have the specified attributes can thus access and finally decrypt the data and also reduce the processing overhead.

In future works, this proposed approach can be implemented and tested on various real-world sensor-cloud problems that are in virtualized WSN environment.

## References

1. Naik, A. K., & Dwivedi, R. K. (2016). A review on use of data mining methods in wireless sensor network. *International Journal of Current Engineering and Scientific Research-IJCESR*, 3(12), 13–20.
2. Dwivedi, R. K., Tiwari, R., Rani, D., & Shadab, S. (2012). Modified reliable energy aware routing protocol for wireless sensor network. *International Journal of Computer Science & Engineering Technology-IJCSET*, 3(4), 114–118.
3. Verma, K., & Dwivedi, R. K. (2016). A review on energy efficient protocols in wireless sensor networks. *International Journal of Current Engineering and Scientific Research-IJCESR*, 3(12), 28–34.
4. Kumar, P., Kumar, R., Kumar, S., & Dwivedi, R. K. (2010). Improved modified reverse AODV protocol. *International Journal of Computer Applications-IJCA*, 12(4), 22–26.
5. Dwivedi, R. K., Sharma, P., & Kumar, R. (2018). Detection and prevention analysis of wormhole attack in wireless sensor network. In *2018 8th IEEE international conference on cloud computing, data science & engineering (Confluence)*, Noida, India (pp. 727–732).
6. Khan, I., Belqasmi, F., Glitho, R., Crespi, N., Morrow, M., & Polakos, P. (2015). Wireless sensor network virtualization: A survey. *IEEE Communications Surveys & Tutorials*, 553–576.
7. Verma, K., & Dwivedi, R. K. (2017). AREDDP: Advance reliable and efficient data dissemination protocol in wireless sensor networks. In *Proceeding of 4th IEEE International Conference on Innovations in Information, Embedded and Communication System-ICIIECS'17*, Tamil Nadu, India (pp. 04–07).
8. Sensor-Cloud. <http://sensorcloud.com/system-overview>.
9. Dwivedi, R. K., Saran, M., & Kumar, R. (2019). A survey on security over sensor-cloud. In *2019 9th IEEE international conference on cloud computing, data science & engineering (Confluence)*, Noida, India (pp. 31–37).
10. Dwivedi, R. K., Singh, S., & Kumar, R. (2019). Integration of wireless sensor networks with cloud: A review. In *2019 9th IEEE international conference on cloud computing, data science & engineering (Confluence)*, Noida, India (pp. 115–120).
11. Dwivedi, R. K., & Kumar, R. (2018). Sensor cloud: Integrating wireless sensor networks with cloud computing. In *5th IEEE Uttar Pradesh section international conference on electrical, electronics and computer engineering (UPCON)*, Gorakhpur, India (pp. 820–825).
12. Zhou, M., Zhang, R., Xie, W., Qian, W., & Zhou, A. (2010). Security and privacy in the cloud: A survey. In *Sixth international conference on semantics knowledge and grid* (pp. 105–112).

13. Alamri, A., Ansari, W. S., Hassan, M. M., Hossain, M. S., Alelaiwi, A., & Hossain, M. A. (2013). A survey on sensor-cloud: Architecture, applications, and approaches. *International Journal of Distributed Sensor Networks*, 917–923.
14. Islam, M. M., Hassan, M. M., Lee, G. W., & Huh, E. N.: A survey on virtualization of wireless sensor networks. *Sensors*, 2175–2207 (2012).
15. Thilakanathan, D., Chenb, S., Nepal, S., Calvo, R., & Alemb, L. (2014). A platform for secure monitoring and sharing of generic health data in the Cloud. *Future Generation Computer Systems*, 102–113.
16. Tu, S., Niu, S., Li, H., Xiao-ming, Y., & Li, M. (2012). Fine-grained access control and revocation for sharing data on clouds. In *26th international parallel and distributed processing symposium workshops and PhD forum* (pp. 2146–2155).
17. Chandrasekaran, B., & Balakrishnan, R. (2016). Efficient pairing computation for attribute based encryption using MBNR for big data in cloud. In *2nd international conference on applied and theoretical computing and communication technology* (pp. 243–247).
18. Li, M., Yu, S., Zheng, Y., Ren, K., & Lou, W. (2013). Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption. *IEEE Transactions on Parallel and Distributed Systems*, 131–143.
19. Pussewalage, H. S. G., Oleshchuk, V. (2016). A patient-centric attribute based access control scheme for secure sharing of personal health records using cloud computing. In *2nd international conference on collaboration and internet computing* (pp. 46–53).
20. Tran, D. H., Nguyen, H. L., & Zha, W. (2011). Towards security in sharing data on cloud based social networks. In *8th international conference on information, communications and signal processing* (pp. 1–5).
21. Hung, N. T., Giang, D. H., Keong, N. W., & Zhu, H. (2012). Cloud-enabled data sharing model. In *IEEE international conference on intelligence and security informatics*.
22. Bethencourt, J., Sahai, A., & Waters, B. (2007). Ciphertext-policy attribute-based encryption. In *Symposium on security and privacy* (pp. 220–239).
23. Sayantani, S. (2015). Secure sensor data management model in a sensor-cloud integration environment. In *Applications and innovations in mobile computing* (pp. 325–332).
24. Sharma, P., & Dwivedi, R. K. (2019). Detection of high transmission power based wormhole attack using received signal strength indicator. In S. Verma, R. Tomar, B. Chaurasia, V. Singh, & J. Abawajy (Eds.), *Communication, Networks and Computing, CNC 2018. Communications in Computer and Information Science* (Vol. 839, pp. 142–152). Singapore: Springer.
25. Dwivedi, R. K., Sharma, P., & Kumar, R. (2018). A scheme for detection of high transmission power based wormhole attack. In *2018 5th IEEE Uttar Pradesh section international conference on electrical, electronics and computer engineering (UPCON)*, Gorakhpur, India (pp. 826–831).
26. Dwivedi, R. K., Pandey, S., & Kumar, R. (2018). A study on machine learning approaches for outlier detection in wireless sensor network. In *2018 8th IEEE international conference on cloud computing, data science & engineering (Confluence)*, Noida, India (pp. 189–192).
27. Yang, Y., & Zhang, Y. (2011). A generic scheme for secure data sharing in cloud. In *40th international conference on parallel processing workshops*.
28. Li, J., Zhao, G., Chen, X., Xie, D., Rong, C., Li, W., et al. (2010). Fine-grained data access control systems with user accountability in cloud computing. In *IEEE second international conference on cloud computing technology and science* (pp. 89–96).
29. Goyal, V., Pandey, O., Sahai, A., & Waters, B.: Attribute-based encryption for fine-grained access control of encrypted data. In *13th ACM conference on computer and communications security* (pp. 89–98 2006).
30. Yang, K., Han, Q., Li, H., Zheng, K., Su, Z., & Shen, X. (2017). An efficient and fine-grained big data access control scheme with privacy-preserving policy. *IEEE Internet of Things Journal*, 563–571.
31. Odelu, V., & Das, A. K. (2016). Design of a new CP-ABE with constant-size secret keys for lightweight devices using elliptic curve cryptography. *Security and Communication Networks*, 4048–4059.

32. Lauter, K. (2004). The advantages of elliptic curve cryptography for wireless security. *IEEE Wireless Communications*, 62–67.
33. Chhabra, A., & Arora, S. (2017). An elliptic curve cryptography based encryption scheme for securing the cloud against eavesdropping attacks. In *International conference on collaboration and internet computing* (pp. 261–269).
34. Gupta, D. S., & Biswas, G. P. (2017). A secure cloud storage using ECC-based homomorphic encryption. In *International journal of information security and privacy* (pp. 550–578).
35. Pérez, S., Rotondi, D., Pedone, D., Straniero, L., Núñez, M. J., & Gigante, F. (2017). Towards the CP-ABE application for privacy-preserving secure data sharing in IoT contexts. In *International conference on innovative mobile and internet services in ubiquitous computing* (pp. 917–926).

# Predictive Model Prototype for the Diagnosis of Breast Cancer Using Big Data Technology



Ankita Sinha, Bhaswati Sahoo, Siddharth Swarup Rautaray  
and Manjusha Pandey

**Abstract** Big data is the collection of thousands of datasets from different application sources just as social media, banking, sales, marketing, etc. In every field, big data technologies are used for analyzing, preprocessing, storing, and generating new patterns for the benefits of the organization. The era of big data technology is nowadays booming [1]. Health care is one of the most important applications of big data. In health care, data exist in different forms like heart rate, blood pressure, blood test, sugar test, cholesterol, and many more. Diagnosis of diseases at an early stage is also very important in healthcare services. Cancer disease is an abnormal cell that negatively affects our body texture and regular functioning body organs. Due to cancer, the death rate is increased as it gets diagnosed at a later stage. Early diagnosis of cancer increases the survival rate of a patient. This paper focuses on the prediction model for the breast cancer diagnosis at an early stage as it increases the chances for successful treatment because of the advanced diagnostics technologies like MRI scans, ductogram, diagnostics mammogram, ultrasound, and many more. So predicting the prognosis of breast cancer increases the survival rate of women. Data mining classification algorithm like SVM, naive Bayes, k-NN, decision tree, etc. combined with analytical tool, which is a promising independent tool for handling huge datasets, is proven better in prediction of the breast cancer diagnosis.

**Keywords** Big data · Health care · Cancer · Breast cancer · Data mining algorithm · Prediction

---

A. Sinha (✉) · B. Sahoo · S. S. Rautaray · M. Pandey  
KIIT Deemed University, Bhuneshwar, India  
e-mail: [ethosankita@gmail.com](mailto:ethosankita@gmail.com)

B. Sahoo  
e-mail: [bhaswati.sahoofcs@kiit.ac.in](mailto:bhaswati.sahoofcs@kiit.ac.in)

S. S. Rautaray  
e-mail: [siddharthfcs@kiit.ac.in](mailto:siddharthfcs@kiit.ac.in)

M. Pandey  
e-mail: [manjushafcs@kiit.ac.in](mailto:manjushafcs@kiit.ac.in)

# 1 Introduction

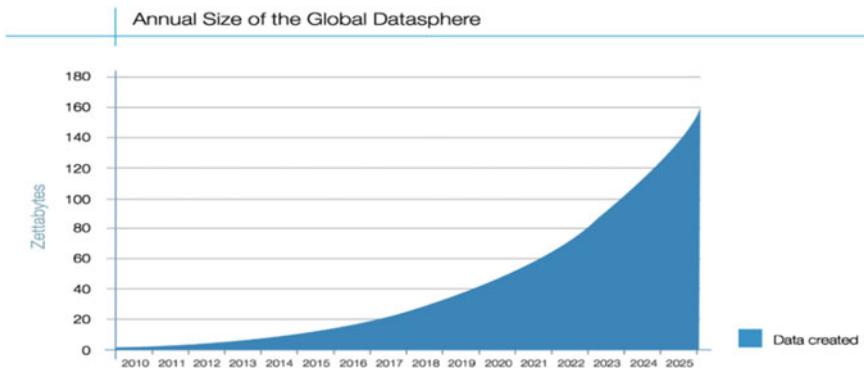
Big data technology is a collection of a huge volume of datasets that cannot be analyzed by traditional computing technologies like batch processing. Not all data are important or useful as it totally depends upon the organization what they do with this large amount of data. Generally, organization does not want to store that much amount of data so they analyze large datasets to discover the hidden pattern, unknown correlations, market trend, and customer preference to make some strategies and also can take good moves for their firm.

The oversized and rapid growth data does not fit in the structures of traditional database formats. There are many difficulties in handling a big amount of data like storage, processing, pattern generation, etc.

Figure 1 shows the volume of big data range of petabyte (PB)–exabyte (EB)–zettabytes (ZB) [2]. The big data technology generated from the client–server interaction generally as recording of customer call or its histories of transaction any many more. Nowadays, big data methodology is getting a lot of importance from organizations for handling huge data and using them in business for its growth. Big data technology examples are data generated from finance or banking sector, Internet hub, smartphone, FM radio frequency identification (RFID), data science, IoT sensors, and streaming are the top seven data main drivers.

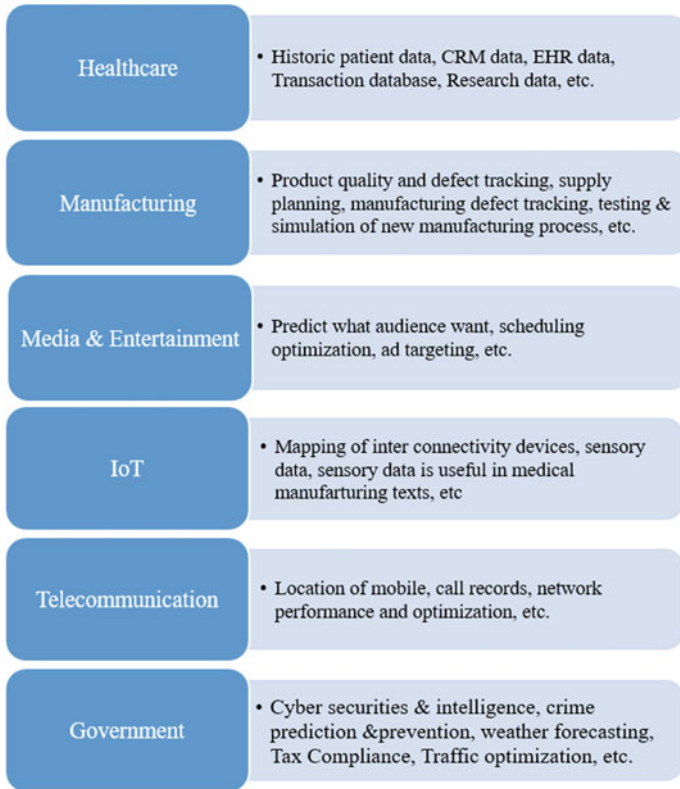
## 1.1 Different Applications of Big Data

According to market strategy, those company organizations that slip the big data opportunities of nowadays will be going to slip the next fortune of “innovation”, “competition”, and “productivity”. For boosting production efficiency and by analysis, it also develops some new data-driven products and related services by using



**Fig. 1** Growth rate of data





**Fig. 2** Different applications of big data

big data tools and technologies which help the companies to make profit. Some of the applications are shown in Fig. 2 [3].

### **1.2 Health Care**

Healthcare service environment is very important and also it is very difficult to maintain high quality, privacy, and affordable cost services. Services like monitoring of patient health, diagnosis, and to inform and educate people about health care and its services are very challenging because of less facilities in government hospitals, private hospital fees, insecure law for patient’s health, and many more. Health care is at top of the iceberg because of its rapid growth and demand of healthcare worker also increased. Health care is essentially driven by various types of information that cover a wide range of topics including mainly administrators, physician assistants, speech therapists, and many more, whose aim is to help people to stay fit and healthy.

Breast cancer is one type of cancer generally found in women's breast area, mostly in ducts and gland. In later section, breast cancer in detail is explained [4].

This paper is organized into sections as follows. Section 2 summarizes the background followed by brief discussion diseases, types of diseases, cancer, stage of cancer, and breast cancer. Breast cancer is discussed in detail like risk factor, various ways of diagnosis, treatment option, and some related work on breast cancer is discussed in Sect. 3. Section 4 is all about the proposed model, i.e., development of the predictive model for the diagnosis of breast cancer using big data technology. Section 5 is the conclusion, which summarizes a brief overview of the proposal followed by future work.

## 2 Background

Disease is a term that does not happen because of any external injuries, and it is an abnormal condition that affects the texture or function of body tissues negatively. The noncommunicable disease is a disease that does not spread while communicating, touching, etc. and some of them are cancer, diabetes, chronic diseases, and many more. Cancer is one of the major diseases that leads to the death of the patient and its diagnosis at a later stage decreases the survival rate.

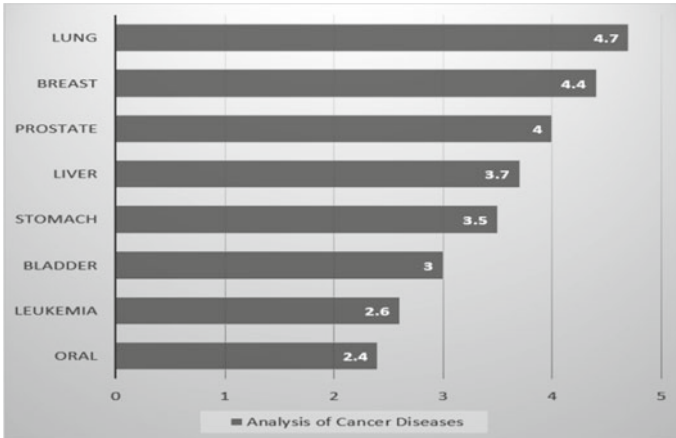
### 2.1 Cancer

Cancer disease is uncontrollable cancerous cell expansion or infectious disease in which anomalous cells get divided into a number of subcells and affect other tissues and in this way, it gets spread throughout the body through blood and lymph. And the movement of the cell from the origin point to throughout the body's tissues is called as metastasis. Cancer is caused due to changes in our regular life schedule like diet change and lack of physical activity, radiation exposure, sun and UV exposure, and genetics causes.

### 2.2 Stages of Cancer

Detection of tumors size and its condition decide to which cancer stage it is classified, helps in diagnosis, and in its treatment.

**Stage 0** Cancer just started and still on its origin point and have not invaded other cells.



**Fig. 3** Analysis of cancer diagnosis

**Stage 1** Cancer is now little bit increased and just started invading other tissues and also spreads to its nearby tissues. It has not spread to lymph nodes and blood vessels or other areas of body.

**Stages 2 and 3** Cancer is now grown up larger and started to invade other tissues and also spread to other tissues by lymph node and blood vessels.

**Stage 4** Disease has spread throughout the body areas. This stage is also called metastatic cancer or advanced cancer.

The effectiveness of treatment is determined by the stages of cancer and its types also. On the basis of a treatment plan, our expert team of doctors determines the stage of the disease by a careful assessment carried out by simple evaluation or advanced diagnostic procedures on a case-to-case basis. There are more than 50 types of cancer like lung, breast, prostate, leukemia, bladder, stomach, skin, oral, and many more.

Figure 3 shows the analysis of cancer diagnosis [5] in the scale of 5. Breast cancer is the second leading cause of death mostly found in women in comparison to men. One in every eight women is facing breast cancer. So the breast cancer diagnosis in an early stage can increase the survival rate. It happens due to changing lifestyle of women like depression, having children after the age of 30, having no children, overweight, etc.

### 3 Breast Cancer and Its Related Work

Breast cancer is a disease in which tumor cell can be present in any area of the breast but mostly found in duct and gland. The breast is combination of tissues like lobules, ducts, fatty and connective, blood vessels, and lymph vessel tissues. Cancer in breast

begins in fatty tissue areas of the breast called as stromal tissues. Also found in surrounding lymph nodes, generally the underarm areas. As the tumor grows in size, it becomes a serious issue and the survival rate decreases.

### 3.1 Risk Factor

Some risk factors of breast cancer are aging (<45), dense breast tissue, hormone disorder in early stage, genetic, defect in BCR1 and BCR2 cell, personal history with breast, physical activity, problem with menstrual cycles, previous chest radiation exposure, benign breast conditions, and oral contraceptive use.

### 3.2 Breast Cancer Diagnosis

When symptoms suggest that cancer is present and for portraying whether the cancer really exists or not, use of diagnostic imaging can help in confirming and if it is present at what stage it is [6]. There are many ways of imaging tests that may be overseen in diagnosis of breast cancer and they are as follows:

**Diagnostic mammograms** are X-rays of the breast in which multiple pictures of doubtful area are taken and then advanced digital mammograms are offered in which multiple pictures are noted down, figured out, stored at the screen so that it can easily be transferred to any hospitals or any physicians.

**MRI scans** use radio waves and a robust magnet in generating detailed information about the pictures of breast's doubtful area. It can be used with a combination of mammograms for the detection of the highly influenced area.

**Breast ultrasound** is mainly used to mark off the stage of cancerous tumors by using sound waves which generally give a depiction of breast tissues or tumor cells.

**Ductogram** test outcome is also an image of the distressed duct where the nipple discharge occurs demonstrating any oddity by implanting contrast medium onto the damaged duct.

### 3.3 Related Works

Some researchers have begun to predict the breast cancer diagnosis or recurrence prediction model with the aim of increasing accuracy by comparing the performance of existing data mining algorithm. This section focuses on research works done in the last 5 years on the prediction of breast cancer and its recurrence and the patient survivability with the missing value. Some of the related works are as follows:

Sakri et al. [7] applied particle swarm optimization (PSO) as feature selection into three renowned classifiers (naive Bayes, K-nearest neighbor, and fast decision tree learner) with an objective to increase the accuracy level of the breast cancer recurrence prediction model by reducing the number of features. Naive Bayes produced better output with and without PSO. Newer algorithm with other feature selection techniques will be included in future.

Alwidian et al. [8] introduce a new pruning and prediction technique based on statistical measures to generate more accurate association rules to enhance the accuracy level of the association classification classifiers and also to solve the problem of estimated measures and prioritization techniques which plays a critical role in rule generations through an efficient weighted classification based on association rules algorithm, i.e., WCBA. They also proposed the use of distinct techniques for weighting, pruning, and prediction with the aim of screening the result on distinct fields in WEKA tool in future.

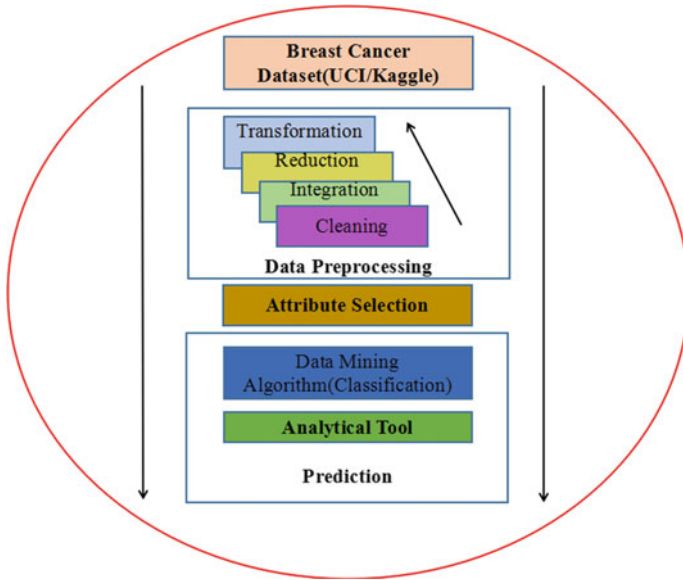
Shukla et al. [9] introduce a robust data analytical model which lifts up understanding of breast cancer survivability in the presence of missing data with better insights into factor associated with patient survivability that also builds cohorts of patients that share identical features by using unsupervised data mining methods, i.e., the self-organizing map (SOM) and density-based spatial clustering of applications with noise (DBSCAN) used to create patient cohort clusters [10]. Decision tree prefers as relevant MLP classifier and the new patient survivability performance, generalized into one of the clusters for better survivability prediction accuracies. As the author used 2D SOM and DBSCAN, so in future, 3D or 4D SOM and DBSCAN will be considered.

Asri et al. [11] introduce Wisconsin breast cancer datasets that are used for comparison of performance in terms of accuracy, precision, sensitivity, and specificity between data mining algorithm (SVM, C4.5, NB, k-NN). The experiments are done on WEKA tool and outcomes confirmed that SVM gives high accuracy, also determined its efficiency in BC prediction with low error rate.

Shah et al. [12] proposed a method in which breast cancer datasets are compared in terms of correctly classified instances, incorrectly classified instances, time taken, kappa statistic, relative absolute error, and root relative squared error between data mining algorithm like decision tree, Bayesian network, and K-nearest neighbor (i.e., random forest, naive Bayes, ibk) on WEKA tool which shows result that naive Bayes gives higher accuracy in lowest time.

## 4 Proposed Model

According to the literature survey, 549 instances were available from 1999 to 2018 but from the patients, diagnosed list from 1999 to 2008 was only 257 instances [13]. The twice increased in instances of breast cancer and the survival rate of corresponding patients diagnosed outcome need good reliable prediction model. Model with standardized huge datasets which fetches important attributes or features is a better



**Fig. 4** Proposed model

possibility of a predictive model. If new data encounters or new attributes are added, it shows the good effective results for the proposed predictive model. There are several attributes for diagnosis of breast cancer which have been described in Sect. 3.3. The proposed model shown in Fig. 4 will take these attributes into consideration and will provide a predictive result [7–9]. The prediction result can be “Begin” and “Malignant” prediction model as shown in Fig. 4 which consists of several steps and these steps consist of UCI/Kaggle-related dataset preprocessing, attribute filtering, data mining algorithm (classification), and prediction modeling using analytical tool and then final prediction as depicted in figure of proposed model.

The proposed model includes the following steps:

**UCI/Kaggle datasets** Dataset will be collected from UCI/Kaggle. The data available here has been collected from xxx hospital. It is available in text format and is converted to csv file format for further processing.

**Preprocessing** of breast cancer datasets requires to extract some meaningful patterns from data and use that information or knowledge. It includes several steps like cleaning which removes noise and inconsistent data. Integration collection of data from multiple sources. Reduction includes data compression, dimensionality reduction. Transformation includes normalization, where data are transformed and consolidated into forms appropriate for mining.

**Attribute selection** is the attribute filtering that considers only important attribute. For example, in the dataset of breast cancer there is no requirement of patient’s job title.

**Data mining learning algorithm** examines the attributes and also enhances attribute relationship in order to generate new knowledge or information. Classification data mining algorithm technique categorized the breast cancer datasets into number of classes (as defined) [14]. The main classification technique is to accurately predict the new data class and also to give effective result for the analysis of huge datasets. For example, naive Bayes, support vector machine, decision tree, k-nearest neighbor, and many more.

**Analytical tool** Datasets are huge in number and for the analysis of large number of datasets we need analytical tool. R programming tool is free software environment for statistical computing and graphics. It is one of the best analytical tools for data preprocessing. Most of the data scientists and researchers use R tool for data analysis. Hadoop is one of the best analytical tools which allows distributed processing of large datasets across multiple clusters and has the ability to handle limitless concurrent tasks. Hadoop computation part is based on MapReduce algorithm which consists of Map and Reduce duty [15]. A duty of Map is to convert the sets of data into another form of datasets and each particular element is broken down into key or value pairs. Reduce duty is to combine all the key or value pairs (output of map) into a smaller element [15, 16].

**Prediction** Predictive model uses statistics to predict the trends and behavior of the model.

## 5 Conclusion and Future Work

The early diagnosis of breast cancer increases the rate of women survival because of advanced diagnosis techniques and the advanced treatment technologies. We have many classification algorithm techniques which can predict and analyze the datasets as it is evolving big data analytical tool. In this research work, we proposed a prediction model which consists of several steps that begin from the collection of datasets from UCI/Kaggle followed by preprocessing of datasets, attribute filtering, and then finally applied data mining classification algorithm like SVM, naive Bayes, k-NN, and decision tree with an analytical tool for the prediction. Future work can focus on the development of the predictive model for the diagnosis of breast cancer using big data technology.

## References

1. Tsai, C. W., Lai, C. F., Chao, H. C., & Vasilakos, A. V. (2015). Big data analytics: A survey. *Journal of Big data*, 2(1), 21.
2. <https://www.zdnet.com/article/by-2025-nearly-30-percent-of-data-generated-will-be-real-time-idc-says/-Growth-rate-of-data>.
3. <https://www.edureka.co/blog/big-data-applications-revolutionizing-various-domains/>.

4. Mohebian, M. R., Marateb, H. R., Mansourian, M., Mañanas, M. A., & Mokarian, F. (2017). A hybrid computer-aided-diagnosis system for prediction of breast cancer recurrence (HPBCR) using optimized ensemble learning. *Computational and Structural Biotechnology Journal*, *15*, 75–85.
5. <http://www2.nau.edu/~gaud/bio372/class/cancer/cancer2.htm>.
6. <https://www.americanoncology.com/diagnosis-of-breast-cancer/>.
7. Sakri, S. B., Rashid, N. B. A., & Zain, Z. M. (2018). Particle swarm optimization feature selection for breast cancer recurrence prediction. *IEEE Access*, *6*, 29637–29647.
8. Alwidian, J., Hammo, B. H., & Obeid, N. (2018). WCBA: Weighted classification based on association rules algorithm for breast cancer disease. *Applied Soft Computing*, *62*, 536–549.
9. Shukla, N., Hagenbuchner, M., Win, K. T., & Yang, J. (2018). Breast cancer data analysis for survivability studies and prediction. *Computer Methods and Programs in Biomedicine*, *155*, 199–208.
10. Park, K., Ali, A., Kim, D., An, Y., Kim, M., & Shin, H. (2013). Robust predictive model for evaluating breast cancer survivability. *Engineering Applications of Artificial Intelligence*, *26*(9), 2194–2205.
11. Asri, H., Mousannif, H., Al Moatassime, H., & Noel, T. (2016). Using machine learning algorithms for breast cancer risk prediction and diagnosis. *Procedia Computer Science*, *83*, 1064–1069.
12. Shah, C., & Jivani, A. G. (2013). Comparison of data mining classification algorithms for breast cancer prediction. In *2013 Fourth international conference on computing, communications and networking technologies (ICCCNT)* (pp. 1–4). IEEE.
13. Jonsdottir, T., Hvanngberg, E. T., Sigurdsson, H., & Sigurdsson, S. (2008). The feasibility of constructing a Predictive Outcome Model for breast cancer using the tools of data mining. *Expert Systems with Applications*, *34*(1), 108–118.
14. Han, J., Pei, J., & Kamber, M. (2011). *Data mining: Concepts and techniques*. Elsevier.
15. Hadoop. [https://www.tutorialspoint.com/hadoop/hadoop\\_mapreduce.htm](https://www.tutorialspoint.com/hadoop/hadoop_mapreduce.htm).
16. Li, D., Park, H. W., Batbaatar, E., Piao, Y., & Ryu, K. H. (2016). Design of health care system for disease detection and prediction on Hadoop using DM techniques. In *Conference on health informatics and medical systems*.



# Recent Dimensions of Data Science: A Survey



Sinkon Nayak, Mahendra Kumar Gourisaria, Manjusha Pandey  
and Siddharth Swarup Rautaray

**Abstract** Nowadays, huge amount of data has been generated and collected in every instance of time. So to analyze them is the toughest task to do. Data are generated and collected in a huge amount from unlike sources such as social media, business transactions, public data, etc. This greater amount of data may be structured, semi-structured, and unstructured one. The data in which analysis is to be performed these days are not only of massive amount but also varies each other by its types, at which speed it is generated and by its value and also varies by different characteristics which is termed as big data. So to examine this vast amount of data and get the relevant information from it, analysis should be done and to analyze this huge amount of data is a greater challenge these days. So to analyze these vast amount of data we need the help of several data analytics tools and methods so that it will be easier to deal with it. This survey paper talks about different tools and techniques used for big data analytics. This survey paper tries to provide a clear idea about the genesis of big data, features of big data, and different tools and techniques used to analyze these huge collections of data.

**Keywords** Big data · Techniques · Clustering · Classification · Architecture

---

S. Nayak (✉) · M. K. Gourisaria · M. Pandey · S. S. Rautaray  
KIIT Deemed to be University, Bhubaneswar, India  
e-mail: [sinkonnayak07@gmail.com](mailto:sinkonnayak07@gmail.com)

M. K. Gourisaria  
e-mail: [mkgourisariafcs@kiit.ac.in](mailto:mkgourisariafcs@kiit.ac.in)

M. Pandey  
e-mail: [manjushafcs@kiit.ac.in](mailto:manjushafcs@kiit.ac.in)

S. S. Rautaray  
e-mail: [siddharthfcs@kiit.ac.in](mailto:siddharthfcs@kiit.ac.in)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_44](https://doi.org/10.1007/978-981-15-0694-9_44)

# 1 Introduction

Aggregation of facts is called data which has some value and size. For any operation to be done, data should be processed into information by the help of knowledge discovery process which is used for finding information which should be logical, novel, useful potentially, and understandable [1]. Data are generated and collected in a huge amount from unlike sources such as social media, business transactions, etc. Big data was characterized by HACE theorem, i.e., heterogeneous, distributed, and centralized control with autonomous sources, survey, and inspect the complex and evolving relationship between data [2]. Big data needs advanced tools and techniques to perform analysis.

## 1.1 Genesis of Big Data

Big data not only simply refer to a huge abstraction of data but it also has some other features on the basis of which we can be able to differentiate between mass of data and big data [3]. Big data becomes enormous expanse of data if one cannot able to process it properly. Huge abstraction of data is generated from various rootages which are listed below:

- **Social media:** Social media like Facebook, LinkedIn, Yahoo, Twitter, Google, etc. hold information about users across the world which is vast in size and to store and process these information is a great challenge.
- **Business transaction:** Business activities just like their sales activities are captured and sometimes these data are engendered at a very reckless speed and millions of records are produced within a fraction of second due to which the scope of the data is huge enough.
- **Public data:** Public data is a kind of data which can be used freely and reused properly without any legal restriction to access these data or to use these data. Public data refers to the data which belongs to a public domain and needs to be updated in a fixed interval of time. For example, in an organization, public data such as job descriptions, press releases, marketing materials, etc. are available to those who are related to the organization and also for them which are external to the organization.

## 1.2 Big Data Characteristics

Feature of Big data is one of the most essential parts for analysis to be done on it. This section of the paper defines the features of big data which consist of 5 V's [4]. Figure 1 shows some of the characteristics of big data such as

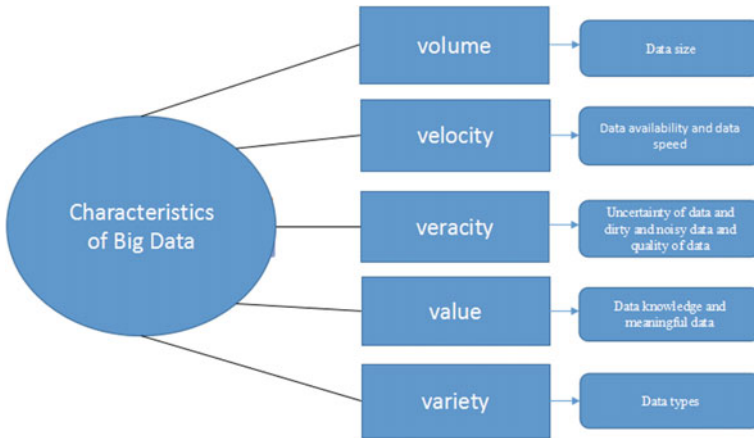


Fig. 1 Big data characteristics

- **Volume:** Volume denotes the giant amount. The term big data itself states the data which consists of bulk amount of data.
- **Velocity:** Velocity refers to the rate of speed of the data at which data are generated and get modified. For example, in social media within a fraction of second greater quantity of videos and photos get uploaded.
- **Veracity:** Veracity states to the features of data in which the giant amount of data which is collected should be 100% correct. There might be a chance that useless data may get included in study, which may affect the correctness of the analytics.
- **Value:** Value is one of the most significant features of big data which measures the effectiveness of data for decision-making drive. For big data, the data size is huge so without any value of the data which is collected it is useless for any operation to be done on that.
- **Variety:** Variety refers to the data which is either in structured format, semi-structured format, or unstructured format. Structured data can be kept either in csv file or Excel sheet. Semi-structured data can be kept in XML file or RSS feed, and similarly unstructured data can be represented in a text format or in graphical like JPEG and GIF formats.

## 2 Big Data Techniques

For scrutinizing the substantial volume of data, it is desirable to extract valuable and concealed information and for performing the analysis, different techniques are used so that it will be convenient for both enterprises and individuals. There are various methods of big data analytics which may belong to different domains.

## ***2.1 Clustering***

Clustering is an unsupervised learning procedure in which groupings of objects are done, where similar kind of entities belongs to same clutch or cluster and dissimilar entities belong to different clutches or clusters. Traditional clustering techniques demand the data in which clustering to be done should be in same format, which is a problem in case of big data [5]. To resolve this problem, clustering technique is divided into two categories, i.e., single machine clustering which resolves the problem of dimensionality and multiple machine clustering which resolves nonconvergent and MapReduce.

## ***2.2 Classification***

Classification comes under supervised learning method in which it predicts a class for each object and assigns them to a target class. Several classification algorithms are developed so that it will work for a distributed environment in a parallel way [5].

## ***2.3 Machine Learning***

Machine learning is a part of data science in which system learns from data, identify patterns, and make decisions with minimum human involvement.

## ***2.4 Representation Learning***

Representation learning comes into picture because of the increase in the dimensionality of the data which can be able to capture a large amount of data and improve the effectiveness in both statistical and computational perspectives [6].

## ***2.5 Deep Learning***

In deep learning, learning can be unsupervised or supervised which automatically learn hierarchical representation. It can be able to capture complex patterns and produce a reasonable output in some cases and is better than human experts [6].

## ***2.6 Parallel and Distributed Learning***

As data sizes are increasing day by day, it is very difficult to store immense quantity of data in a single machine. Distributed and parallel learning is the learning where data are stored across multiple machines and can be accessed in a parallel way so that it can achieve faster training [6]. Distributed and parallel learning helps to scale up the learning algorithm which is different from classical learning algorithm [7].

## ***2.7 Transfer Learning***

Transfer learning is different from other learning techniques because in other learning techniques the training and testing data have same feature space but in case of transfer learning the data which is to be trained and tested have different feature space. Transfer learning has an advantage of learning from the past and applies these to solve the new ones [6].

## ***2.8 Active Learning***

Learning from a bulk amount of data is the most critical task to do. Active learning helps to do the learning when data are in huge amount and not labeled [6]. It labeled some samples so that time and cost will be reduced in the process of learning. Active learning used to attain maximum accuracy by the help of labeled sample which tends to reduce the cost.

## ***2.9 Kernel-Based Learning***

Kernel-based learning is a very almighty learning which helps to increment computational capableness [6]. It is a nonlinear learning algorithm in which samples are mapped to infinite-dimensional space from their original space which is done by inner product operator with a suitable kernel function.

## ***2.10 Regression Analysis***

Regression analysis is used for finding association among the variables like among independent variables and a dependent variable. It finds out the association among the variables which are concealed by randomness.

### 2.11 Hashing

Hashing is used for transforming a string of characters to a fixed length of values which represents the same value as original. By the use of hashing, speedy reading and writing can be done and query speed will also expand but the problem is to find out a suitable hash function by which these things can be done.

### 2.12 Indexing

Indexing is a method used for diminishing the disk access time so that the time will be a smaller amount while doing insertion, deletion, and update in occasion of both structured data in case of relational database and numerous techniques used in occasion of semi-structured and unstructured data. But the difficulty we face is while indexing it needs an additional space to store the index file which upsurges the cost.

Figure 2 describes different clustering techniques, classification techniques, and tools used for analyzing big data.

## 3 Clustering

Clustering is a learning procedure in which data objects are divided into groups in such a manner that the data objects which have similar kind of entities are grouped into same clutch and the data objects having different kind of entities are grouped into

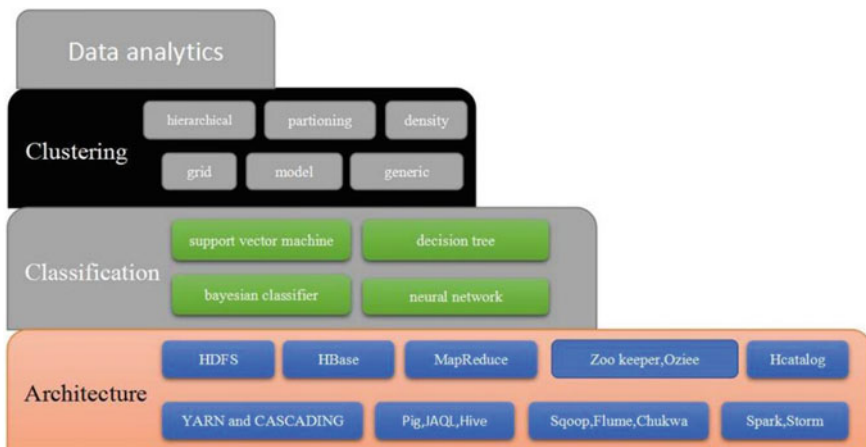


Fig. 2 Different clustering, classification, and tools for data analytics

a different clutch. There are various types of clustering techniques and depending upon various factors they are classified as follows.

### ***3.1 Partitioning Based***

In partitioning-based clustering algorithm, the data entities are divided in such a way that every clutch must have at least one entity and each entity must fit into only one clutch. The dividers are called clusters. One needs to specify the number of clusters primarily [8, 9].

### ***3.2 Hierarchical Based***

In hierarchical clustering algorithm, the data constituent is delineated in a hierarchical based which can be either agglomerative or divisive. In agglomerative clustering, initially each item is in each own cluster and then clusters are merged with each other iteratively. In divisive clustering, initially all items are in one cluster and then clusters are divided such that each item consists of its own cluster [8, 9].

### ***3.3 Density Based***

In density-based clustering algorithm, data items are divided depending upon their densities. Clusters of random shapes can be revealed and also can handle outliers by density-based clustering [8, 9].

### ***3.4 Grid Based***

In grid-based clustering algorithm, the data objects are divided into grids. It scans the dataset once to compute the grids and employ a grid which can collect the data and clustering algorithm is applied to it instead of the dataset which leads the computation time to be faster [8].

### **3.5 *Model Based***

In model-based clustering algorithm, clusters are determined automatically based on standard statistics taking outliers into consideration, so model-based clustering algorithm is more robust [8].

### **3.6 *Generic Based***

Generic-based clustering algorithm is an unsupervised clustering algorithm. In this algorithm, it searches for the input data which occur frequently and group them into a cluster [9].

## **4 Classification**

Classification belongs to supervised learning method in which it predicts a class for each object and assigns them to a target class. Classification aims to predict the target class for each data in a dataset accurately. There are several classification techniques available to analyze bulk amount of data which are given below.

### **4.1 *Support Vector Machine***

It is based on decision planes on decision margins. A decision plane can be defined as the one which splits up between a bunch of objects and belongs to unlike class. SVM is a supervised learning model which is used mutually for categorization or classification as well as regression analysis [10].

### **4.2 *Decision Tree***

In decision tree, the data constituents are basically depicted in a tree-like structure which comprises of a source node, branch nodes, terminal nodes, and branches in which branch nodes denote attribute, terminal nodes denote class labels, and branches denote the outcome [5].



### ***4.3 Bayesian Classifier***

Bayesian classifier is a probabilistic method to solve the classification problem. Bayesian classification is a supervised learning where categorization or classification is cooked based on Bayes theorem [5].

### ***4.4 Neural Network***

Neural network for classification uses multilayer perceptron. Neural network is used to cluster by using a set of rules which is done basically by the help of decision tree [11].

## **5 Architecture**

To process using traditional techniques is very difficult due to complexity and volume of big data. On the basis of various perspective of big data, i.e., data storage, data processing, data querying, data accessing, and data management different tools are used.

### ***5.1 HDFS***

HDFS is the major storage system of Hadoop which has the capability to store a huge amount of data and process them as well. HDFS architecture is a master–slave architecture in which Master Node keeps track of the data to be stored in the cluster and manages the Data Nodes, whereas Data Nodes are responsible for read–write operations on the data blocks which are present inside them. If Data Nodes fail to keep the replica of block then Name Node creates another replica of that block [11].

### ***5.2 HBase***

On Hadoop, HBase is an open-source, non-relational, and distributed database, which can able to do read–write operations in a random way and also in an efficient amount of time. HBase data are stored in row on table rather than column [11].

### **5.3 *MapReduce***

MapReduce is a programming model which can able to process an immense magnitude of data through some applications. The mapper function is responsible for mapping the subtask to different nodes and reducer function is responsible for reducing the responses from the nodes to a single one [12].

### **5.4 *YARN and CASCADING***

YARN stands for Yet Another Resource Negotiator which is a distributed storage system in Hadoop whose job is cluster management, job scheduling. CASCADING is a high-level Java API which helps to hide the complexities in MapReduce programming. It is an application development platform for data applications on Hadoop [2].

### **5.5 *Pig, JAQL, Hive***

Pig architecture is used for analyzing the large data in Hadoop. It is also responsible for ETL process, i.e., extraction, transformation, and load, which uses a language called Pig Latin, a high-level language. So users are depending upon Pig Latin to do any data operation. JAQL is used for JSON query processing for big data which is a data processing language. Hive is database which is used for processing structured and semi-structured data. It is not similar to relational database but supports some parts of structured query language, semi-structured data [2].

### **5.6 *Sqoop, Flume, Chukwa***

Sqoop is a data transfer tool which is used to transfer a huge amount of data between Hadoop and relational database. It also loads data directly into Hbase or into Hive and imports data from sql query directly. Flume helps to collect, assemble, and transfer bulk amount of data from distributed resources. Chukwa acts just like a detector to the distributed system, and for data collection it takes the help of HDFS and for analysis it takes the help of MapReduce [2].

## 5.7 *Spark, Storm*

Spark is used for increasing the computational power of Hadoop which was developed by Apache software foundation. It can be used in either way for storing and for processing. Storm is an open-source parallel computing system which uses ZooKeeper. It does not run on Hadoop but can do read–write operations on HDFS [12].

## 5.8 *Hcatalog*

Hcatalog is a metadata service provider which also provides an interface of read–write operation on MapReduce and Pig. It is a data processing tool for Hadoop to write data [2].

## 5.9 *Zookeeper, Oziee*

Zookeeper works as a centralized monitoring server in Hbase which takes care of configuration information and maintains coordination between distributed applications. Oziee is an open-source service provided by Apache Hadoop which consists of such programs such that all jobs are done in their desired order in Hadoop environment [2].

This section of the paper describes about the tools and techniques used for analyzing big data. To analyze a huge collection of data which are different from each other by several characteristics is a toughest task to do nowadays.

## 6 Conclusion

This survey paper tries to give an idea about what big data is and the sources from which enormous amount of data are generated. It also designates the features of big data by which one can differentiate big data and data which are of huge quantity. Any operation to be done on bulk amount of data such as store data, access these stored data, modify these data whenever it needed, and delete the data is the biggest challenge. As the data size is increasing day by day, examining such substantial volume of data is a serious issue. So to examine these data the usage of various techniques and tools are required, which ease the analysis to be done.

To analyze these bulk data, we need to have a clear idea about the tools and techniques. This paper describes the tools and techniques used for analyzing big data such as Hadoop, MapReduce, different clustering techniques, different classification techniques, different machine learning techniques, etc. This survey paper overviews

the techniques which are proposed how to deal with a massive amount of data and protect those data so that no one will access or modify it except the authentic users.

## References

1. Patil, S., et al. (2017). Data analysis & classification methodology for knowledge discovery in big data. In *2017 international conference on inventive systems and control (ICISC)*. IEEE.
2. Dhyani, B., & Barthwal, A. (2014). Big data analytics using Hadoop. *International Journal of Computer Applications*, 108(12).
3. Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, 19(2), 171–209.
4. Ertekin, S., & Hopper, G. (2006). Efficient support vector learning for large datasets. *Grace Hopper Celebration of Women in Computing*.
5. Wu, X., et al. (2014). Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering*, 26(1), 97–107.
6. Qiu, J., et al. (2016). A survey of machine learning for big data processing. *EURASIP Journal on Advances in Signal Processing*, 1, 67.
7. Peteiro-Barral, D., & Guijarro-Berdiñas, B. (2013). A survey of methods for distributed machine learning. *Progress in Artificial Intelligence*, 2(1), 1–11.
8. Fahad, A., et al. (2014). A survey of clustering algorithms for big data: Taxonomy and empirical analysis. *IEEE Transactions on Emerging Topics in Computing*, 2(3), 267–279.
9. Arora, S., & Chana, I. (2014). A survey of clustering techniques for big data analysis. In *2014 5th international conference confluence the next generation information technology summit (Confluence)*. IEEE.
10. Saravanan, K., & Sasithra, S. (2014). Review on classification based on artificial neural networks. *International Journal of Ambient Systems and Applications (IJASA)*, 2(4), 11–18.
11. Saraladevi, B., et al. (2015). Big data and hadoop—a study in security perspective. *Procedia Computer Science*, 50, 596–601.
12. Tsai, C.-W., et al. (2015). Big data analytics: A survey. *Journal of Big Data*, 2(1), 21.
13. Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META Group Research Note*, 6(70).
14. Acharjya, D. P., & Ahmed, K. (2016). A survey on big data analytics: Challenges, open research issues and tools. *International Journal of Advanced Computer Science and Applications*, 7(2), 1–11.

# Sentiment Analysis: Usage of Text and Emoji for Expressing Sentiments



Shelley Gupta, Archana Singh and Jayanthi Ranjan

**Abstract** Social media generates massive volume of unstructured data by means of blogs, social networking sites, purchasing sites, etc. This huge amount of data plays a very crucial role in determining the opinions and sentiments of people toward a product, services, popularity of artists, celebrities, etc. The accurate and meaningful analysis of this online media further helps in developing product quality, making strategies, and decision for a company or a personality. The online media consists of both text and emoji. Emoji is pictorial representation of facial expression, objects, weather, animals, etc. In this paper, the web document of various worldwide famous male and female personalities has been examined to determine the amount of usage of emoji and presence of text and emoji both in expressing sentiments. This paper also presents the comparative analysis of the total usage of emoji among the most followed male and female personalities in Twitter. The major application of our paper is that it exhibits that the sentiment analysis is more accurate and complete when done for both text and emoji.

**Keywords** Sentiment analysis · Text and emoji usage · Twitter dataset analysis · Most twitter followers

---

S. Gupta (✉)

Department of Information Technology, ABES Engineering College, Ghaziabad, Uttar Pradesh, India

e-mail: [shelley.gupta@abes.ac.in](mailto:shelley.gupta@abes.ac.in); [shelley.g17@gmail.com](mailto:shelley.g17@gmail.com)

A. Singh

Department of Information Technology, Amity School of Engineering and Technology, Noida, India

e-mail: [asingh27@amity.edu](mailto:asingh27@amity.edu)

J. Ranjan

Department of Information Technology, Institute of Management Technology, Ghaziabad, India  
e-mail: [jranjan@imt.edu](mailto:jranjan@imt.edu)

## 1 Introduction

The tremendous increase in the online shopping trends and social networking websites has led to the requirement of analyzing the customer's emotions and trends of purchasing with sentiment analysis tool, facilitating the brand to develop product quality, adjust market strategy and customer service from the user-generated messages.

The user messages consist of both text and emoji. Emojis are the pictorial representation of user sentiments, facial expressions, places, sports, and weather. They have provided the world with a new language to express their emotions in colorful, appealing, and amusing way, with the need of few or no word in the messages [1].

Since late 2006, online text analysis has become a focused research area by means of social media content, blogs, forums, etc. With the advancement of technology and social media platforms like Facebook, Twitter, forums, etc. user has the choice of expressing their emotions with text and emojis both. In this paper, tweets as the web documents of various famous personalities are downloaded, preprocessed, and analyzed. The main objectives of the paper are as follows:

- To analyze and evaluate the web document of various worldwide famous male and female personalities.
- To evaluate the count of usage of emoji by different male and female personalities, individually.
- To evaluate the total usage of text and emoji both in the web document of different male and female personalities, individually.
- To evaluate the count of usage of emoji among male and female class.
- To evaluate the total usage count of text and emoji both in male and female class.
- To perform the comparative analysis of the average percentage of emoji usage among males and females in their web documents.
- To perform the comparative analysis of the average percentage of documents with text and emoji both in the web documents of males and females.
- To exhibit that complete and accurate sentiment analysis can be done with approaches involving both text and emoji due to huge availability of online content with both text and emoji.

## 2 Literature Review

Sentiment analysis has become an active research area of Natural Language Processing Language at many levels of granularity [2]. Sentiment analysis can be classified at three main levels: sentence level [2], document level [3], and aspect level [4]. Document-level sentiment analysis determines the positive, negative, or neutral opinion a document is expressing, taking the complete document as a single unit. Sentence-level sentiment analysis determines whether the sentence is subjective or

objective. If the sentence is subjective it determines whether the sentence is expressing a negative or positive opinion. Aspect-level sentiment analysis determines all the aspects of the opinions a document is expressing.

The major techniques of machine learning used in sentiment analysis are based on supervised learning and unsupervised learning. In supervised learning, the trained dataset is needed whereas in unsupervised learning labeled dataset is not needed.

The base of sentiment analysis is words classified as positive, negative, and neutral sentiments. Most of the research in sentiment analysis is English text based using lexicon corpora. General Inquirer [5] is a lexicon-based approach that used pointers to link similar meaning words. SentiWordNet 3.0 [6] is a lexicon-based approach and an enhancement of WordNet [7], associated with three numerical scores, positive, negative, and neutral taking synsets as base. The SenticNet [8] works on the approach of concept-level opinion mining, consisting of 14,244 commonsense concept.

Earlier emoticons were used in expression to express the emotions with text on web. Emoticons are also the facial representation using characters like punctuation marks, numbers, and letters. In the work [9], the data has been trained based on the language and emoticons. Nowadays, the emoticons have been replaced by emojis. The emojis were first included in Japan mobile phones in 1997 [1] and operating system of Apple in 2011. The Unicode Standard has incorporated some emoji character set from 2010. The new Emoji Version 11.0 has been released in 2018, with which total number of approved emoji has reached a total number of 2,823 [10, 11].

In [11], the sentiments of emojis were calculated from the sentiment of tweets. In [12], Arabic tweets were used to analyze which emoji is used most frequently and classify them as anger, disgust, joy, and sadness. Most of the research in sentiment analysis has been done with text [13] only or emoji [1] only but not involving both.

The objective of the paper is to determine that the usage of both emoji and text is huge in social media for expressing social sentiments and exhibiting the great importance of sentiment analysis approach involving text and emoji both by analyzing the tweets of most followed male and female [14] personalities on Twitter.

### 3 Methodology

The most followed Twitter celebrities [14] are divided into two categories, i.e., male and female.

- The web documents of each personality are analyzed to determine the following:
  - The total count of emoji used.
  - The total count of web document consisting of both text and emoji.
- The web documents of each category are analyzed to determine the following:
  - Average count of emoji in web documents of both categories.
  - Average count of web documents with text and emoji in both categories.

- The comparison of
  - Average percentage of total emoji used by both categories.
  - Average percentage of documents with text and emoji in both categories.

Most of the research in the field of sentiment analysis is done with text only or emoji only. Therefore, we aimed at exhibiting that a sentiment analysis approach will give more accurate and appropriate results if it considers both text and emoji.

## 4 Proposed Approach

Sentiment analysis is a popular research area which plays an important role in analyzing the sentiments of people posted by them by means of various social networking sites, e-commerce sites, forums, blogs, etc. The steps of approach are discussed in Fig. 1:

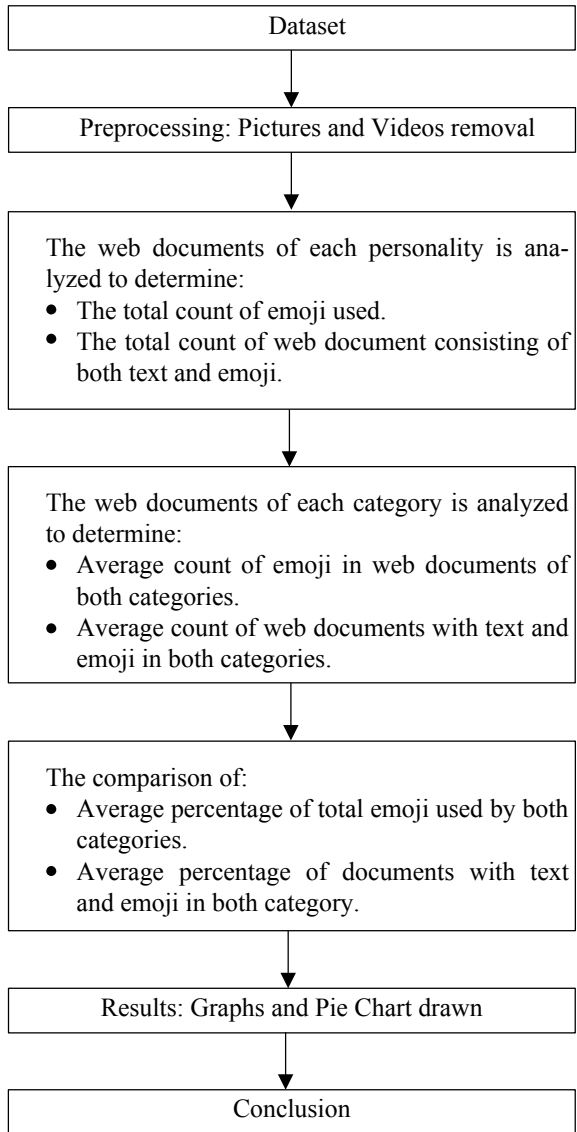
- Step 1: The data collection approach is discussed in Fig. 2. The tweets posted on twitter consist of text and emoji both in combination. Thus, the tweets of most followed personalities on twitter [11] are downloaded, classified as male and female. In this paper, 26 male and 26 female Twitter handles are collected. Broadly, the last 200 tweets of personalities are provided by the Twitter API.
- Step 2: The last 200 tweets of each personality are preprocessed to remove videos and pictures.
- Step 3: Tweets of each personality of both class are analyzed to determine total count of emoji used and total count of web document consisting of both text and emoji.
- Step 4: The tweets of each category are analyzed to determine the average count of emoji and average count of web tweets with both text and emoji in both categories.
- Step 5: The comparison of percentage of total emoji used by both categories and the percentage of tweets with text and emoji in both categories.
- Step 6: The bar graph is drawn for each of the 26 male and 26 female personalities, exhibiting
  - Total tweet considered for each personality.
  - The total count of emoji used by each personality in the total tweets considered.
  - The count of tweets consisting of both text and emoji together out of total tweets considered.

The pie charts were drawn showing the following analysis:

- The average percentage of emoji count in both classes.
- The average percentage of tweets consisting of text and emoji both out of the total tweets considered in both classes.



Fig. 1 Proposed approach



Step 7: Based on statistical analysis we determine the class:

- Using more emoji.
- Consisting of more tweets with both text and emoji.

On the basis of usage of emoji among users in tweets decides the role of an approach doing sentiment analysis considering both text and emoji.

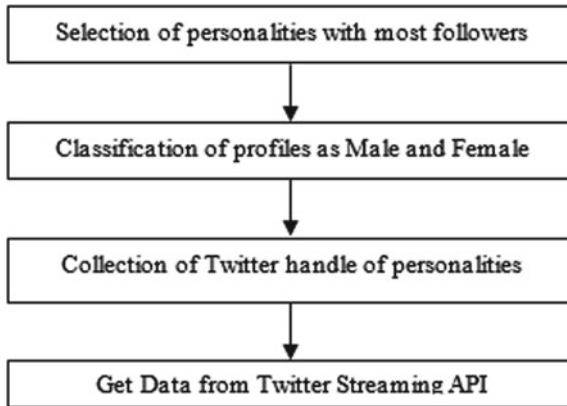


Fig. 2 Data collection approach

### 5 Implementation and Results

The dataset of tweets has been downloaded by using Twitter API implemented in Python [15]. A Twitter application is created using an existing Twitter account. The module of Python called emoji is also used to include emoji in dataset. Some of Python library used involves *Tweepy* for connecting Twitter Streaming API, *Openpyxl* to read and write in excel, *Pathlib* to work with file paths, *re*, *workbook*, etc.

The two bar graphs, *Male Tweet Analysis* Fig. 3 and *Female Tweet Analysis* Fig. 4, were obtained.

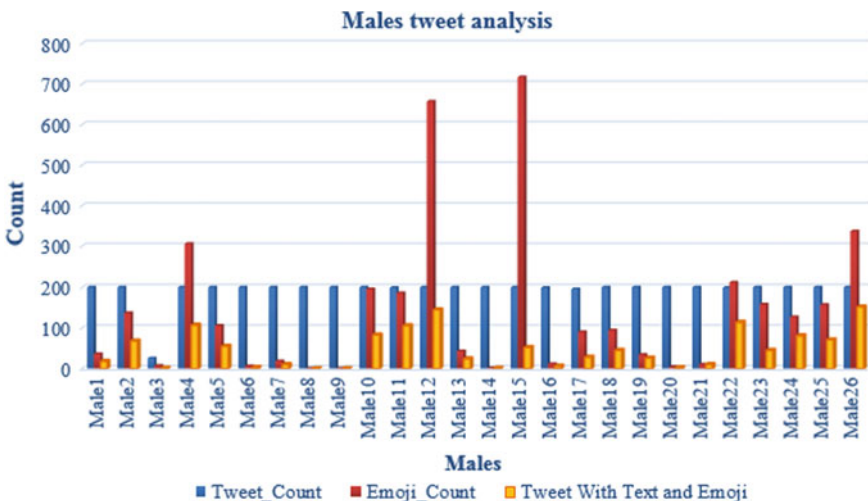


Fig. 3 Male tweet analysis

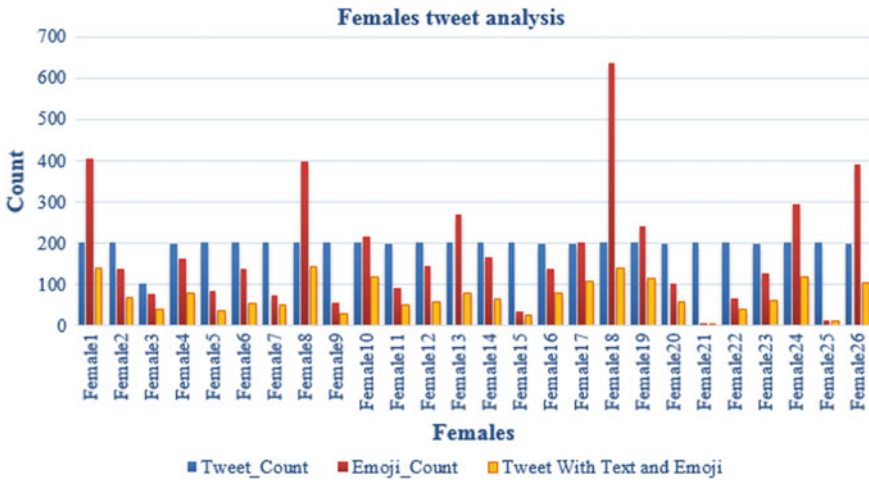


Fig. 4 Female tweet analysis

- *Male Tweet Analysis*

The male tweet analysis graph analyzes each of the 26 male personalities individually. The count of emoji used by most of the males is more than or almost equal to the total count of tweets considered. For example, male 4, male 10, male 11, male 12, male 15, male 22, male 2, and male 26. The tweets with both text and emoji are also significant in number.

- *Female Tweet Analysis*

The female tweet analysis graph analyzes each of the 26 female personalities individually. The count of emoji used by most of the females is more than or almost equal to the total count of tweets considered. For example, female 1, female 4, female 8, female 10, female 13, female 17, female 18, female 19, female 24, and female 26. The tweets with both text and emoji are also significant in number.

Table 1 shows results obtained by further analysis done on the two bar graphs: *Male tweet Analysis* and *Female tweet Analysis*.

**Table 1** Male and female tweet analysis

Class analysis	Male	Female
Total number of tweets considered broadly for each of 26 members of each class, i.e., male and female	200	200
Average count of emoji used	140	179
Average tweet count with both text and emoji	48	71
Average percentage of emoji used	44	56
Average percentage of tweet count with text and emoji both	40	60

The detailed analysis of the two pie carts is given below:

- *Analysis of male and female total emoji count:*  
The percentages of average emoji usage by male and female are 56% and 44%, respectively. Thus, usage of emoji is more in female than in male Fig. 5.
- *Analysis of male and female tweet count with both text and emoji:*  
The percentages of average tweets with both text and emoji are 60% and 40%, respectively, Fig. 6. The usage of tweets with text and emoji both is more in females as compared to males.

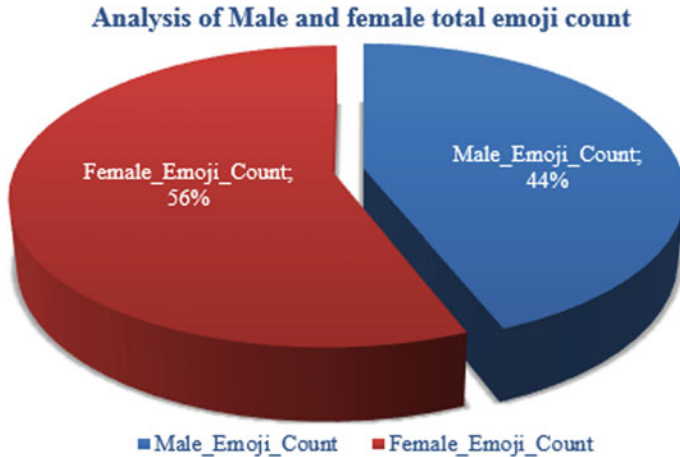


Fig. 5 Analysis of male and female total emoji count

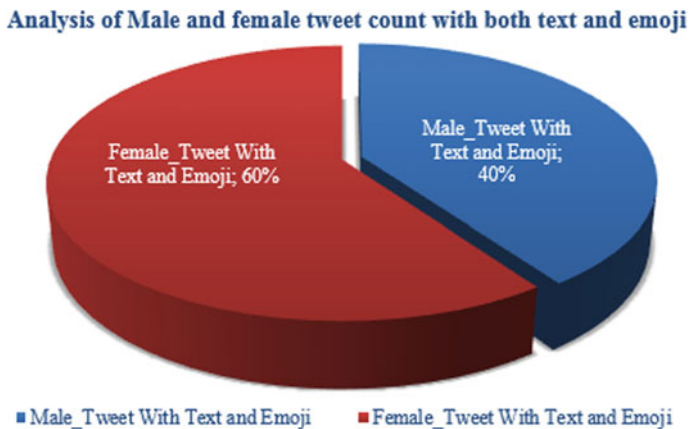


Fig. 6 Analysis of male and female tweet count with text and emoji both

The results show that the availability of emoji is much in online web documents. A sentiment analysis approach will be complete and accurate if it considers both text and emoji, whereas most of the research involves text only or emoji only.

## 6 Conclusion and Future Work

Our approach analyzes the total count of emoji in tweets and count of tweets with text and emoji is significant in number. Also the usage of emoji among females is more than the males across the world. In this paper, we exhibit the usage of emoji is tremendous among the people across the world to express their sentiments on social media. Thus, the requirement is to have sentiment analysis approaches that consider both text and emoji. As our results are promising, we envisage several directions for future work. Our findings are based on the two classes, i.e., male and female. We would like to expand our finding for many other important classifications as well.

## References

1. Fernández-Gavilanes, M., Juncal-Martínez, J., García-Méndez, S., Costa-Montenegro, E., & González-Castaño, F. (2018). Creating emoji lexica from unsupervised sentiment analysis of their descriptions. *Expert Systems with Applications*, *103*, 74–91.
2. Liu, B. (2012). *Sentiment analysis and opinion mining*. San Rafael, Calif: Morgan & Claypool.
3. Moraes, R., Valiati, J., & Gavião Neto, W. (2013). Document-level sentiment classification: An empirical comparison between SVM and ANN. *Expert Systems with Applications*, *40*, 621–633.
4. Zhang, W., Xu, H., & Wan, W. (2012). Weakness finder: Find product weakness from Chinese reviews by using aspects based sentiment analysis. *Expert Systems with Applications*, *39*, 10283–10291.
5. Stone, P., Bales, R., Namenwirth, J., & Ogilvie, D. (2007). The general inquirer: A computer system for content analysis and retrieval based on the sentence as a unit of information. *Behavioral Science*, *7*, 484–498.
6. Baccianella, S., Esuli, A., & Sebastiani, F. (2010). SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *Proceedings of LREC* (Vol. 10).
7. Miller, G. (1995). WordNet: A lexical database for English. *Communications of the ACM*, *38*, 39–41.
8. Cambria, E., Havasi, C., & Hussain, A. (2012). SenticNet 2: A semantic and affective resource for opinion mining and sentiment analysis. In *Proceedings of the 25th International Florida Artificial Intelligence Research Society Conference, FLAIRS-25* (pp. 202–207).
9. Read, J. (2005). Using emoticons to reduce dependency in machine learning techniques for sentiment classification. In *Proceedings of the ACL Student Research Workshop*. ACL.
10. New Emojis in the 2018 Emoji List. <https://blog.emojipedia.org/157-new-emojis-in-the-final-2018-emoji-list/>.
11. Kralj Novak, P., Smailović, J., Sluban, B., & Mozetič, I. (2015). Sentiment of emojis. *PLoS One*, *10*, e0144296.
12. Hussien, W., Tashtoush, Y., Al-Ayyoub, M., & Al-Kabi, M. (2016). Are emoticons good enough to train emotion classifiers of arabic tweets?.

13. Tripathy, A., Agrawal, A., & Rath, S. (2016). Classification of sentiment reviews using n-gram machine learning approach. *Expert Systems with Applications*, 57, 117–126.
14. Find out who's not following you back on Twitter, Tumblr, & Pinterest. <https://friendorfollow.com/twitter/most-followers/>.
15. Standard search API. <https://developer.twitter.com/en/docs/tweets/search/api-reference/get-search-tweets.html>.

# Factors Affecting Psychological State of Youth in India



Jagreeti Kaur, Archana Singh, Sumit Kumar and Sunil Kumar

**Abstract** The best of brain working does not mean the absence of some sort of mental illness. It is beyond that. The ability to deal with the pressure of negative situation varies greatly from one person to another. The fast-paced lifestyle influenced humans' communication and thinking power. This paper identifies the significant reasons behind this contemporary but devastating lifestyle. Some emotional state of mind like being loved, self-esteem, confidence, and breakups leads to adopting difficult behavior like smoking, drinking, and usage of various drugs. To examine this problem among youth, the data was collected from a reputed university's students those who were hosteler or non-hosteler and using factor analysis, it was found that major contribution in the drinking, smoking, and drugs was social status, loneliness, and depression.

**Keywords** State of mind · Youth · Factors · Socioeconomical · Lifestyle

## 1 Introduction

In today's life, after the revolution of technology and the effect of that in our life, the mental health is still one of the biggest problems in human's life, especially for today's youth. Everyone has their own mental health problems in their lives. In youth generation, same things happened due to lots of pressure, sadness, happiness, and depression they face with these problems. The mental illness depends on the

---

J. Kaur (✉)

ABES Engineering College, Ghaziabad, India  
e-mail: [jagreeti.kaur@abes.ac.in](mailto:jagreeti.kaur@abes.ac.in)

A. Singh · S. Kumar · S. Kumar  
Amity University, Noida, Uttar Pradesh, India  
e-mail: [archana.elina@gmail.com](mailto:archana.elina@gmail.com)

S. Kumar  
e-mail: [skumar59@amity.edu](mailto:skumar59@amity.edu)

S. Kumar  
e-mail: [skumar58@amity.edu](mailto:skumar58@amity.edu)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_46](https://doi.org/10.1007/978-981-15-0694-9_46)

different reasons and issues, and today they face these issues due to the reason of not sharing their problems with families and friends. This is one of the main concerns that today's young generations have. The lack of sharing of problems leads to short-term solutions which can act as an asymptotic relief to the problems. The youngsters do not choose to share problems may be due to lack of trust, experience, and lifestyle. The community, peer pressure, to perform best in all things, and social status have been few findings in the recent research. According to United Nations (UN), India is the world largest youth population with 365 million of 18–30 years old. So, all these issues like lack of communication, habits, suicides, stress of studies, and depression that today's youth generations are facing and the attempt to find out the reasons for all the above-said issues were the major concern. Much research has been done to study youth's psychological and behavioral state of mind in US, Canada, and some Europe countries. The lacuna of intensive study about youth erratic behavior in India is the motivation of this paper.

In this paper, the study was carried out to study the psychological, behavioral, and medical factors affecting the youth who were drawn toward some drugs, alcohol, and smoking habits are explored. The data was analyzed for the youth of age between 18 and 30 years. The data was analyzed using factor analysis and resulted in identifying significant factors affecting psychological state of youth in India.

The paper is organized as follows. In the following section, youth mental issues are explored. In Sect. 2, the related work done is discussed; in Sect. 3, research methodology is discussed followed by experiments and results in Sect. 4, and in Sect. 5 conclusion and future scope is mentioned.

## ***1.1 About Mental Health***

Mental health is including the social well-being and emotions of us. Mental health is the way of our behavior, thinking, and how we act, feel, and show our attitude to others. These are the things that express our mental health. Also mental health shows how we can go or specify our daily life stress and other problems that we face in our life. Whenever we face any mental health issues, it directly affects our thinking, attitude behavior, and the way of daily life. Here are some factors which show whether we are having any mental health issues:

- Feeling hopeless and helpless.
- Sleeping too much or too less, eating.
- Family history of emotional wellness issues.
- Using drug, drinking, and smoking.
- Life experience like abuse or trauma.
- Biological factors like brain chemistry or genes.

All these factors are the signs for facing mental health problem that we experience in our life. In our life, we will see or experience different problems or face them.



## ***1.2 Types of Mental Health***

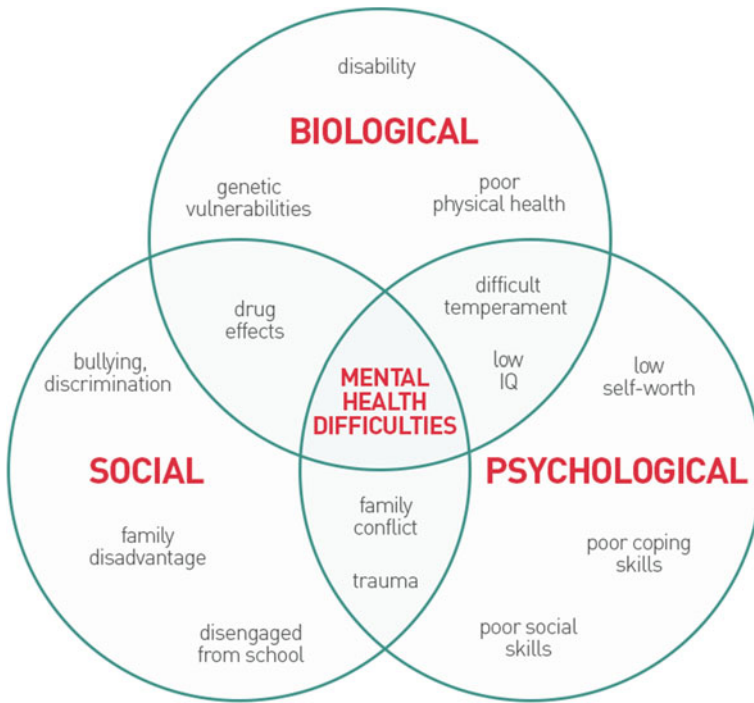
In the world of mental health issues, there are too many types and categories and there are some conditions that are known as mental issues or illnesses, such as

- **Anxiety disorder:** The people who have this illness have fear and dread when they face a situation. And also they show some physical sign from themselves like their body starting to sweat and will have rapid heart beating. This illness is diagnosed when the person does not respond specifically or appropriately and does not have any control to the response.
- **Fluctuating mood disorder** is persistent feeling of much happy and sadness.
- **Eating disorder:** This illness is due to evolving attitudes and emotions. In this illness, the patient will not have control on their eating orders and behaviors involving the weight and the food.
- **Personality disorders:** In this illness, the attitude and the personality of the person will change and it will completely differ from others. The person will have more depression and will create lots of issues in daily life.

## ***1.3 Youth Mental Health Problem***

Youth mental health disorder is when a young person who has a negative impact and has difficulties on his/her thinking, mode or behavior at home, school, or among the friends and peers. For diagnosing the mental health disorder, health professionals look at grouping of symptoms, for example, if a young person has a mental disorder problem for long time of being sad or down and little interest or pleasure in things they would normally enjoy. A mental health professional will consider that these behaviors or the symptoms may lead to the diagnosis that a young person is having “clinical depression”. There is no single cause for youth or young person to experience a mental health, and generally there are multiple causes and factors that play effect on our mental. Numbers of models explain the mental health issues or disorders, and our understanding changes over the time with advancement in research, especially in the brain sciences. The main purpose of diagnosis is to enable and to aware them to make informed. Recommendations should be done to the young person who has mental illness and what support or treatment may be useful for them. One of the explanations is that mental illness is caused by social, biological, and psychological factors that interact in a variety of ways. All refer to a biopsychological model of mental issues (Fig. 1).

- Social or life events (for example, peer or school stressors, family).
- Biological (for example, inherited vulnerabilities).
- Psychological (for example, poor adapting aptitudes).



**Fig. 1** Factors leading to mental health difficulties [7]

## 2 Related Work Done

The extant research examined the rate of contacting mental health care and primary care professionals by individuals, before people commit suicide. The research reviewed 40 studies for which there was information regarding the rates of health care contact and examined age and gender differences among the subject. It is not known that what degree contact with primary care and mental health provider can prevent suicide [1, 2]. The United States is actively discussing the growing movement for developing and expanding mental health programs [3, 4]. So these programs represent partnerships between schools and community mental health agencies to expand the range of mental health services provided by schools. The work has been carried upon the formulation of the strategy for improving the youth mental health programs, which includes the analysis of existing psychological health assessment program in a community. The authors also outline steps to avoid negative attitudes that may arise among professionals from different disciplines when they collaborate to expand school-based programs [5]. The development of more comprehensive school-based program will bring people together who have not had any close working relationships and also the collaboration of people from different systems for the betterment of youth. The state of depression and stress can be minimized with

the community of stakeholder, mental health, and education professionals working together to solve problems, plan, and share knowledge with the goal of developing a system for the youths and families in their community. In the past 2 years than in the past three decades, the mental health needs of youth in the juvenile justice system [6] have received more attention at the federal level combined. A number of factors contributed to this change. They include increasing reliance on the justice system to care for individuals with mental illnesses, growing recognition of mental health need of youth in general [7, 8]. In various countries, lots of work have been induced on how to improve the understanding of the behavior analysis of youth, time spent on social media, and peer pressure for the betterment of the mental health in youngsters [9, 10]. The research carried out lacks to address the whole number of issues pertaining the effect on the psychological state of the youth in India. This paper focusses on psychological, behavioral, and biological factors affecting the youth of India.

### 3 Proposed Model

In the proposed feature model as shown in Fig. 2, by studying the social behavior, psychological responses in various situations and biological, and if any medical health in record were all considered. The usage of drugs, alcohol, and sedatives in youth could be due to various factors. The significant factors like loneliness, time pass, to show off to their peer friends, excessive usage of social media, disturbed

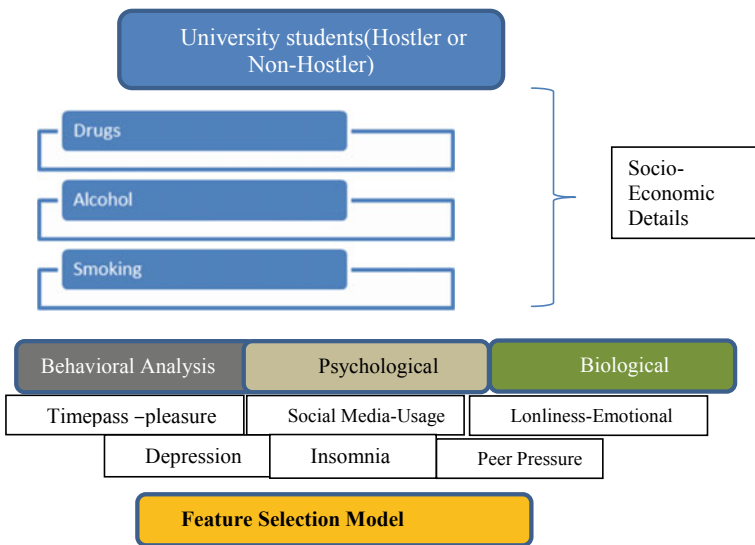


Fig. 2 Proposed feature selection model

sleep patterns, and depressing mental state affect the overall holistic development and psychological state of youth.

In the proposed model, the socioeconomic conditions are taken into the input as a significant parameter. Using dimension reduction algorithm in SPSS, significant factors were found that affect the psychological state of the youngsters.

The significant factors leading to the above problems and lifestyle in youth were found by applying factor analysis method using SPSS statistical package. The factor analysis is one of the techniques used for dimension reduction. The objective of this method is to spot structure through information dimension reduction and to try and to do information reduction, essentially to seek out correlations between the variables and respondents.

## 4 Experiments and Results

The data was collected and to collect the data, the questionnaire was prepared with the help of extant research and advices of experts, talking, and brainstorming sessions with youth of age between 18 and 30 years including types of questions like frequency of using social media, reasons of using social media, smoking, reason for smoking, reason for drinking alcohol, usage of drug, reason for usage drug, usage of antidepression medicine, and usage of sleeping pills. Within these questions we can easily find the interests of today' youth and how they pass their times. After collecting the data, we changed the data into Likert scale from (1–5) 5–Extremely important, 1–Not important. The data was collected based on the problems and issues related to youth lifestyle and psychological state of mind. The sample size was about 200 dataset. It resulted that 52% of youth are purposeless and perform various tasks just for unknown fun, 30% have the problems related to insomnia, and 25% have depression. The correlations matrix was generated to find out the most correlated attributes. It resulted in 44% factors significant at the level of 0.01. By taking the 12 relevant parameters from the questionnaire structure form, feature extraction/factor analysis is done. To assess the overall significance of the correlation matrix with the KMO and Barlett's test, overall significance is at the 0.0001 level which is 106.984, depicting nonzero correlations. By using KMO and Barlett's test, the value of Chi-Square test comes out to be 106.984 with degree of freedom 45 and it is showing significance between attributes.

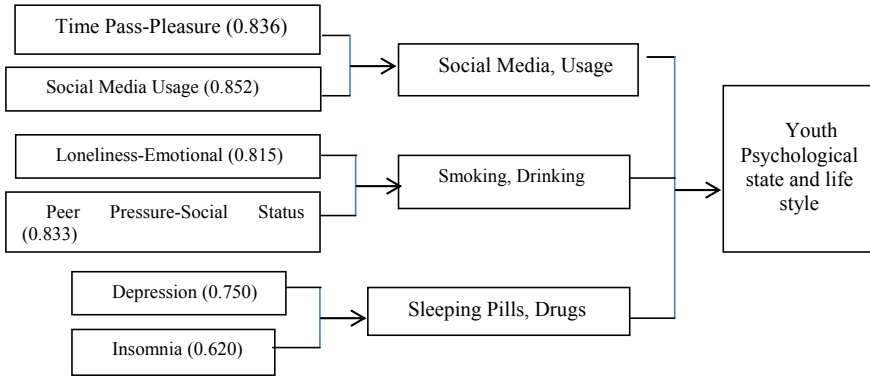
Table 1 shows the factors that were extracted. Under the column of "Rotation Sums of Squared Loadings", it shows all those factors which have eigenvalue >1 as the condition was set in the beginning. The values under the column "% of Variance" depict how much of the total variability (cumulatively variables are taken) can be accounted for by which of these summary scales of factors. Factor 1 shows 13.311% of variability in all 10 variables. Factor 2 shows 12.040% of variability out of 10 and so on.

The above model in Fig. 3 shows that the respondents depict that the usage of social media time pass pleasure factor as extremely important (52%, 51%) and very

**Table 1** Extraction method: principal component analysis

Component	Initial eigenvalues		Extraction sums of squared loadings		Rotation sums of squared loadings	
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	1.331	13.311	13.311	1.331	13.311	13.311
2	1.204	12.040	25.351	1.204	12.040	25.351
3	1.155	11.549	36.900	1.155	11.549	36.900
4	1.057	10.566	47.466	1.057	10.566	47.466
5	1.039	10.393	57.859	1.039	10.393	57.859
6	0.949	9.491	67.350			
7	0.920	9.204	76.553			
8	0.852	8.520	85.073			
9	0.756	7.564	92.637			
10	0.736	7.363	100.000			

Extraction method: Principal component analysis



**Fig. 3** Model representing significant influences affecting psychological state in youth

important and important (55%), loneliness-Emotional implications (47%) and peer pressure for maintaining the social status (52%) are significant factors for smoking and drinking, respectively. Depression and insomnia (58%, 54%) lead to the usage of sleeping pills and drugs.

## 5 Conclusion and Future Scope

Today’s youth dwell in the pool of technology and escapism from emotional outbursts. The usage of the technology is not adequate or normal. The excessive usage of technology and lifestyle of drinking, smoking, and usage of drugs and sleeping pills collectively affect the health and psychological state of young people who are the future leaders. This paper catches the significant reasons behind this contemporary but devastating lifestyle. The data was collected from university students those who were hosteler or non-hosteler and it was found that major contribution in the drinking, smoking, and drugs was social status, loneliness, and depression.

The work can be extended by applying data mining techniques to generate various significant rules and statistical structural equation model to expose the hidden constructs and behavior patterns in youth.

**Acknowledgements** We would like to acknowledge the efforts of our student Khalid Nazari (M.Tech (IT), Amity University, Noida) who helped us in collecting the data from his hostel to study about the lifestyle of students and the administration who allowed us to carry out this research. This data and work are true to the best of our knowledge.

## References

1. Austin, J., Johnson, K. D., & Gregoriou, M. (2010). *Juveniles in adult prisons and jails: A national assessment (NCJ 182503)*. Washington, DC: Bureau of Justice Assistance.
2. Luoma, J. B., Martin, C. E., & Pearson, J. L. (2002, June). Contact with mental health and primary care providers before suicide: A review of the evidence.
3. Brown, G. W., Harris, T. O., & Bifulco, A. (1986). Long-term effects of early loss of parent. In M. Rutter, C. E. Izard, & P. B. Read (Eds.), *Depression in young people* (pp. 251–296). New York: Guilford Press.
4. <http://www.physio-pedia.com>.
5. Jane Costello, E., Mustillo, S., Erkanli, A., Keeler, G., & Angold, A. (2003, August). MRCPSych. Prevalence and development of psychiatric disorder in childhood and adolescence, *60*.
6. Cocozza, J. J., & Skowrya, K. R. (2000, April). Youth with mental health disorders: Issues and emerging responses, *7*(1), 3–13.
7. Prins, P. J., & Hanewald, G. J. (1999). Coping self-talk and cognitive interference in anxious children. *Journal of Consulting and Clinical Psychology*, *67*(3), 435–439.
8. Treadwell, K. R. H., & Kendall, P. C. (1996). Self-talk in youth with anxiety disorders: States of mind, content specificity, and treatment outcome. *Journal of Consulting and Clinical Psychology*, *64*(5), 941–950.
9. Abram, K. M., Teplin, L. A., McClelland, G. M., & Dulcan, M. K. (2003, November). Comorbid psychiatric disorders in youth in juvenile detention, *60*.
10. Weist, M. D., Lowie, J. A., Flaherty, L. T., & Pruitt, D. (2001, October). Collaboration among the education, mental health, and public health systems to promote youth mental health, *52*(10).

# Enhancing Personalized Response to Product Queries Using Product Reviews Incorporating Semantic Information



Payal Aich, Manju Venugopalan and Deepa Gupta

**Abstract** E-commerce is very much in trend in the current era for selling/purchasing products online. Hence, consumers tend to visit question answer forums to know about a product before making a purchase. The proposed work is to build a web application which would form an alternative to a traditional question answer system. Rather than focusing on a knowledge-based question answer system, the proposed work attempts to mine reviews related to the product and provides critical reviews on the product which are relevant to the question asked by the consumer. This application would be used by any user looking for supporting critical reviews related to product functionalities. Given a question answer dataset and a review dataset on a product, the similarity between the questions and the reviews is calculated and three top reviews which are most relevant to the question along with their relevance score are the output of the system. The model uses powerful similarity measures based on WordNet and Word embedding in addition to the basic similarity measures based on cosine similarity and TF-IDF. The model is evaluated in terms of how well the sentiment extracted from the output reviews of the proposed model confines with that of the answer in the question answer dataset.

**Keywords** Similarity measures · Word to vector · Word embedding · Q/A system

## 1 Introduction

Text mining deals with the analysis of huge volumes of text data so as to derive useful insights. Text mining can be helpful for an organization, when it derives

---

P. Aich (✉) · M. Venugopalan (✉) · D. Gupta (✉)  
Department of Computer Science and Engineering, Amrita School of Engineering, Amrita  
Vishwa Vidyapeetham, Bengaluru, India  
e-mail: [p.aich493@gmail.com](mailto:p.aich493@gmail.com)

M. Venugopalan  
e-mail: [v\\_manju@blr.amrita.edu](mailto:v_manju@blr.amrita.edu)

D. Gupta  
e-mail: [g\\_deepa@blr.amrita.edu](mailto:g_deepa@blr.amrita.edu)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_47](https://doi.org/10.1007/978-981-15-0694-9_47)



valuable information from text-based content such as word documents, email, and data streams from social networks like Facebook, Twitter, and LinkedIn. Natural language processing (NLP) techniques in combination with statistical modeling are often employed in text mining. It is really very challenging to perform mining on unstructured data because natural language text is often inconsistent or ambiguous, inconsistent semantics and syntax, sarcasm, slang usage, and varying trends on social media are some of the factors which cause the major challenges. Text analytics is still considered as an emerging technology, with a wide range of applications which include spam detection [1], business intelligence [2], social media analysis [3, 4], sentiment analysis [5–8], and so on. The immense amount of data available online in the form of text is the foremost reason for the growing popularity of text mining. Users share reviews about products, movies, travel experiences, etc. online. These reviews provide the first level of feedback when considering the purchase of products online. Prospective purchasers tend to have a lot of queries about the product which could either be generic in nature or specific to their requirements. Such questions can be answered either through a manual search across the reviews or otherwise put forward the question in a Q/A system. But the disadvantage with a traditional Q/A system is that the response time would be more if the question does not exist in the knowledge base and the user needs to wait until someone responds to his question. Moreover, the answer would reflect only the responder's perspective. Consumer reviews are a valuable source of data in this context in the sense that there is a high probability of arriving at more sensible answers to questions posed and it reflects multiple user opinions. Hence, the proposed work explores the possibility of combining question answer systems and product reviews, and hence design a system which would respond to a query about a product in terms of a set of reviews most relevant to the query.

Section 2 explores a few prominent research works on question answer system and product review systems, Sect. 3 discusses the proposed methodology in detail, Sect. 4 discusses the datasets used for the experimentations, Sect. 5 explains the methodology used for model evaluation, and Sect. 6 presents the experimental setup and results. Finally, the conclusions and the scope for further enhancements are discussed in Sect. 7.

## 2 Related Work

The huge volume of data available online has given immense opportunities for researchers to explore various dimensions in text mining. This section has tried to explore the prominent works done in the design of question answer systems and the works which have experimented product reviews. An exploration of various approaches experimented in the field of QA systems has been presented in [9, 10]. A case-based reasoning method (CBR) [11] has been explored for the design of an intelligent question answering system. The methodology claims easier access to the knowledge base compared to the rule-based traditional systems. The results of similar problems solved prior aids in the faster processing of new questions. Rather

than improving the accuracy the focus of the model was to improve the response time. Antonio proposed a novel approach [12] for integrating question answering and database querying systems. The idea of combining these systems is to mutually benefit each other but the challenge is in that while question answer systems deal with natural language, query systems are meant for structured query languages. The researcher was able to initiate a new direction but concluded that there were more challenges which included information representation and reasoning capabilities on such a representation. A QA system [13] based on an information retrieval (IR) methodology and a validation step for removing incorrect answers has been another attempt in this field. The IR module could extract additional information from the analysis of questions, which contributed in a positive way but the validation module because of too strict constraints removed more correct answers. A trial to incorporating semantic measures for a general open-domain question answering system has been attempted in [14]. The main objective of this research was to improve name-entity-based QA systems which fail with common-noun-based questions. Their results proved that the incorporation of semantic features like WordNet contributed to the model. Product reviews are another major source of data that has been explored to a large extent to extract user sentiments. A notable exploration [15] is to mine user-generated product reviews available online by using a novel approach using the features that can be mined from user-generated reviews along sentiment expressions that are associated with these features. Another approach is mining of textual reviews [16] to summarize product experiences that can serve as recommendation systems by combining aspects of product similarity and feature sentiment. McAuley [17, 18] and team proposed a novel idea of using product reviews to formulate answers for queries asked about a product formulating this as a machine learning problem using a mixture-of-experts-type framework where each review is considered an expert. The survey reveals that QA systems have improved in terms of performance and beyond that researchers have tried to give new dimensions to the domain. The essence of the proposed approach is derived from McAuley's work borrowing only the idea of combining a QA dataset and a review dataset which is based on an expectation-maximization (EM) method to model ambiguity and subjectivity in product-related opinion question answering systems.

The proposed approach is a QA model to answer questions posed on products by extracting the most relevant reviews corresponding to the question by using powerful linguistic similarity measures. The power of the proposed model lies in the incorporation of powerful semantic measures. Word embedding and WordNet are used in addition to statistical measures to calculate the similarity for each question-review pair.

**Table 1** Information on datasets

Dataset	No. of questions and answers	No. of reviews
Beauty products	171	7500
Automotive	150	5090
Baby products	140	9723

```
{
  "reviewerID": "APYOBQE6M18AA",
  "asin": "0615391206",
  "reviewText": "I've enjoyed using this set in my kitchen for over a decade,
  so much so that I purchased it as a house-warming gift for
  my son. Great way to make perfect rice every time (or other
  grains), without having to employ a stand-alone rice
  cooker."}

{
  "QuestionID": "Q3AB4912"
  'asin': '0615391206',
  'questionType': 'yes/no',
  'questionText': 'does the steamer basket have a handle so you can easily
  lift it out when hot?'
  'answerText': 'Yes, it has a heat resistant handle'}

```

**Fig. 1** A sample product review and question answer

### 3 Datasets

The experiments have been carried out on question answer and review datasets borrowed from Amazon.<sup>1</sup> Using their unique product IDs, the Q/A datasets and review datasets are mapped. The proposed approach is confined to questions belonging to the category of binary questions, which have a Yes/No inclination as reflected by the “Question Type” field. Datasets corresponding to 20 product IDs extracted the statistics which are presented in Table 1. A sample of product reviews and question answer dataset is given in Fig. 1 where *asin* indicates the *product id* used to map them.

### 4 Proposed Methodology

The proposed approach for mining reviews in response to queries is delineated in Fig. 2. The sequence of steps includes preprocessing of both questions and reviews,

<sup>1</sup><http://jmcauley.ucsd.edu/data/amazon/>, <http://jmcauley.ucsd.edu/data/amazon/qa/>.

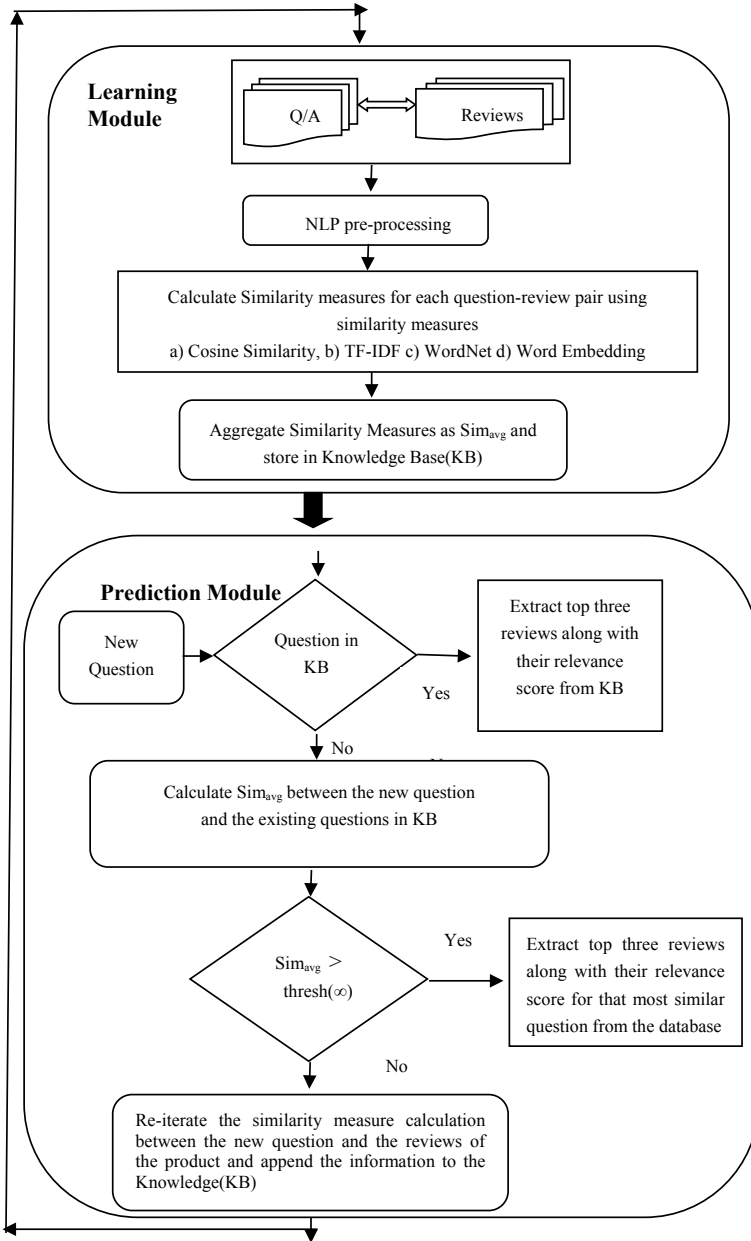


Fig. 2 Proposed methodology

extracting similarity measures, aggregating the similarity measures to find the relevance score for each question–review pair, and hence finding the most relevant reviews for every query.

#### 4.1 Data Preprocessing Module

The process of converting raw data into representable features is known as preprocessing. The unstructured training data which includes both queries and reviews are subjected to the preprocessing steps, tokenization, stop word removal, and lemmatization. Tokenization is the process of splitting the text into smaller units corresponding to the words of the language considered. In NLP, these units are referred to as Tokens. The sentence “*Can this Bluetooth speaker help for small parties*” will be tokenized into a list of tokens as  $\{Can, this, Bluetooth, speaker, help, for, small, parties\}$ . Stop words are words like in, if, for, etc. which do not convey any meaning, and hence are filtered out. Stop word removal of the tokenized sentence would output  $\{Can, Bluetooth, speaker, help, small, parties\}$ . Lemmatization reduces the word to the base form. The purpose of performing lemmatization is to categorize different inflectional forms of a word under the same group. The output of the input sentence after going through the lemmatization phase would be  $\{Can, Bluetooth, speaker, help, small, party\}$ .

#### 4.2 Extraction of Similarity Measures

After the preprocessing stage, the reviews and the questions are in the form of a set of tokens. The similarity between a question/query related to a particular product and a set of reviews about the same product is calculated using different similarity measures as listed in the following subsections.

**Cosine similarity.** It is a simple similarity measure that operates on bag-of-words representations of the review and the query. The unigrams in the set consisting of all questions  $Q$  and all reviews  $R$  corresponding to a product together form the bag of words. The similarity between a question  $q \in Q$  and a review  $r \in R$  is calculated as given in Eq. (1), where  $q$  and  $r$  are the vector representation of a question and a review corresponding to the same product based on the frequency-based representation of bag of words of model and  $\|q\|$  and  $\|r\|$  are the magnitudes of the vectors. The metric is well preferred for text data because of its low complexity for sparse data.

$$\forall q \in Q \text{ and } \forall r \in R \quad \text{Sim}_{\cos}(q, r) = \frac{q \cdot r}{q \|r\|} \quad (1)$$

**TF-IDF-based similarity.** TF-IDF stands for term frequency–inverse document frequency, and is a weighted measure often used in information retrieval and text

mining. This statistical measure resolves an issue with measures like the cosine similarity whereby common but irrelevant words can dominate the ranking function. Variations of the TF-IDF weighting scheme are often used to determine a document's relevance given a user query. The proposed work has implemented the OkapiBM25 metric [19] based on the probabilistic retrieval framework which is given in Eq. (2).

$$\text{Sim}_{\text{TF-IDF}}(q, r) = \sum_{i=1}^p \frac{\text{IDF}(q_i) f(q_i, r) (k_1 + 1)}{f(q_i, r) + k_1 \left(1 - b + \frac{b|r|}{\text{avgrl}}\right)} \quad (2)$$

where  $q_i$  is the  $i$ th unique word in the question, the total number of unique terms being  $p$ ,  $f(q_i, r)$  is the number of occurrence of the  $i$ th unique word of question in the current review,  $k_1$  and  $b$  are constants designated a value 0.5,  $|r|$  is the length of current review, and  $\text{avgrl}$  is the average length of reviews, both calculated in terms of the number of tokens after preprocessing.  $\text{IDF}(q_i)$ , the inverse document frequency, is calculated as given by Eq. (3) where  $N$  is the total numbers of reviews,  $n(q_i)$  is the total number of reviews in which the  $i$ th word in  $q$  has occurred, and the constant  $k$  is assigned a value 0.5.

$$\text{IDF}(q_i) = \frac{N - n(q_i) + k}{n(q_i) + k} \quad (3)$$

**WordNet-based similarity.** Cosine- and TF-IDF-based similarities are based on bag-of-words model which perform a word-to-word mapping for measuring similarity. But such measures fail to capture semantic relatedness which cannot be measured beyond word similarity. Knowledge-based measures like WordNet quantify semantic relatedness of words using a semantic network. WordNet [20] is a lexical database for English language which groups words into sets of synonyms called synsets and records a number of relations among these synonym sets or their members. The similarity between two words can be extracted from WordNet which is actually a measure extracted from the synsets of both the words. Using this measure the similarity between a question and a review is calculated based on [21] as given in Eq. (4) where  $\text{sim}_{\max}(w, r)$  is the maximum value in the similarities of the word  $w$  in the question to each word in the review  $r$ ,  $\text{sim}_{\max}(w, q)$  is the maximum value in the similarities of the word  $w$  in the review to each word in the question  $q$ , and  $\text{IDF}(w)$  is calculated as similar in Eq. (3).

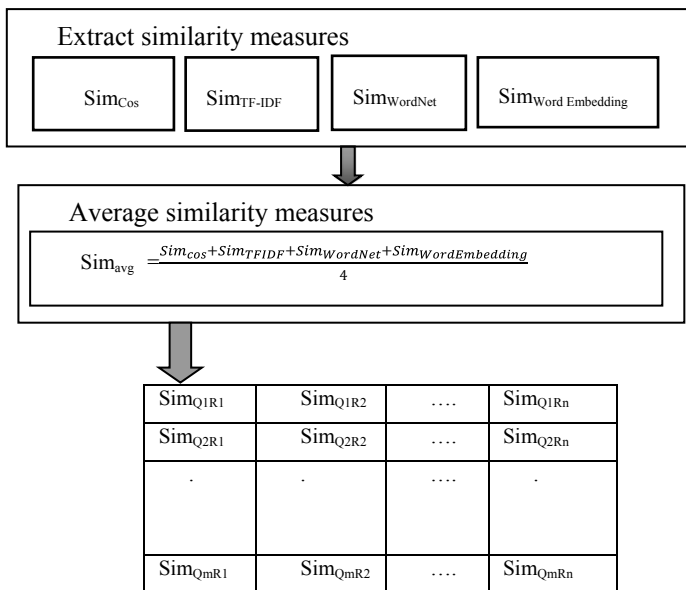
$$\text{Sim}_{\text{wordnet}}(q, r) = 0.5 \frac{\sum_{w \in q} \text{sim}_{\max}(w, r) * \text{IDF}(w)}{\sum_{w \in q} \text{IDF}(w)} + \frac{\sum_{w \in r} \text{sim}_{\max}(w, q) * \text{IDF}(w)}{\sum_{w \in r} \text{IDF}(w)} \quad (4)$$

**Embedding-based similarity.** Word embedding represents a set of language modeling features in natural language processing (NLP) where words from the

vocabulary are mapped to vectors of real numbers. Word embedding is a type of word representation that allows words with similar meaning to have a similar representation. A distributed representation for text has arrived at that is perhaps one of the key breakthroughs for the impressive performance of deep learning methods on challenging natural language processing problems. The low-dimensional vector representation for words is obtained by training a neural network on a large corpus to predict a word from the given context. The context is a window of surrounding words. Using word embedding model a vector representation is created for every word. The proposed model has borrowed pretrained vectors from Glove’s [22] implementation which had trained vectors from Wikipedia 2014 Corpora. The word2vec [23] model had been implemented to arrive at a similarity score between any two words. These scores are in turn used to arrive at a similarity measure  $Sim_{word-embedding}$  for each question–review pair using Eq. (3) itself the only difference that similarity measures used here are those derived from the word embedding model.

### 4.3 Aggregation of Similarity Measures

Using the four similarity measures based on cosine similarity, TF-IDF based, WorNet, and Word Embedding, the proposed approach arrives at four scalar values depicting the similarity between a question and a review. The average of these four similarity scores is computed for each question review pair as depicted in Fig. 3. This score



**Fig. 3** Averaging similarity measures

implies the relevance of the review with respect to the question, and hence referred to as the relevance score. This information gets stored in a database.

#### 4.4 Prediction Models

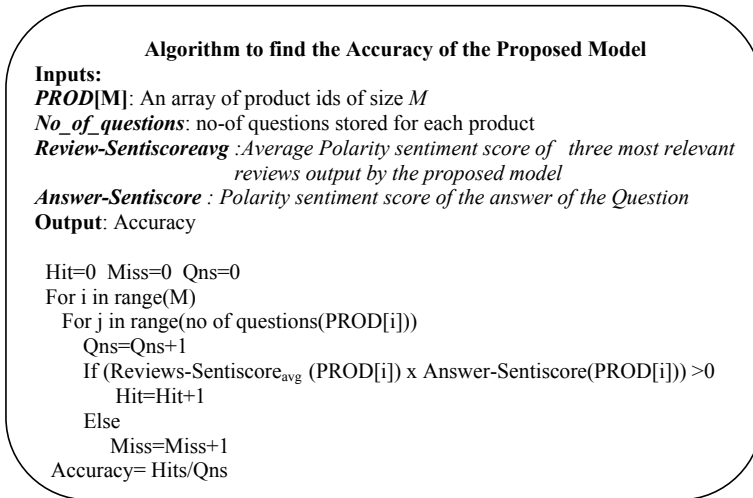
When a new question is posed by a user in a Q/A environment, the model will check if it matches any existing question in the database or is entirely new. If it is an existing question in the database then the top three reviews along with their relevance score will be extracted from the database and will be the output of the model. If it is a non-existing question in the database, then the question is compared with all the existing questions for that product in the database by calculating an aggregated similarity measure for each question–question pair using the same four similarity measures. If the similarity score between the new question and the most similar question is greater than a threshold  $\alpha$  then the top three reviews and their relevance scores are used as the output. But if it is an entirely new question or in other words if there are not any existing questions with similarity measure of at least  $\alpha$ , then the new question gets appended to the database as a new question and it passes through all the phases in the methodology to arrive at the top three reviews and the relevance score corresponding to the new question.

### 5 Model Evaluation

Our model has confined its experimentations to questions which have been labeled *Yes/No* in the Amazon dataset. As clear from Fig. 1, these categories of questions expect answers with inclination toward a *Yes/No* result probably with a justification of their personal experience.

The evaluation of the model has been done in terms of how efficiently the sentiment extracted from the reviews output by the model matches the sentiment expressed from the answer of the question. For this task, the polarity score of the top three reviews that has been output by the model is calculated using the TextBlob package in Python and their average score is determined as *Review-Sentiscore<sub>avg</sub>*. The polarity score of the answer for that particular question from the Q/A dataset is similarly calculated as *Answer-Sentiscore*. If *Review-Sentiscore<sub>avg</sub>* matches the *Answer-Sentiscore* in polarity orientation, it supports the model performance and is treated as hit, else is a miss. Thus, the accuracy of the model is calculated in terms of how many hits the model achieves against the total number of questions. A pseudocode for model evaluation is depicted in Fig. 4.





**Fig. 4** Pseudocode for model evaluation

## 6 Experimental Results and Analysis

The entire framework of the proposed model has been implemented in Python 3. NLTK package in Python has been used for the preprocessing steps. For model evaluation, the sentiment score has been calculated using the Text Blob package which is a Python library. Text Blob goes along the piece of text, finding words and phrases it can assign polarity to, and it averages them all together for the longer text and thus assigns a sentiment score to the answers of the question and the top three reviews output by the model. The evaluation is done on all the three datasets and the accuracy of all the three datasets is reported in Table 2. The designed GUI provides the interface for the user to choose product and enter a query and the model displays the most relevant queries and their relevant scores. The GUI has for each product id a list of previous query asked and an empty field for the user to enter a query of his choice (Fig. 5).

**Table 2** Results of model evaluation

Datasets	Accuracy (%)
Beauty product	92.98
Automotive product	90.31
Baby product	92.64



Fig. 5 GUI for the proposed model

## 7 Conclusion and Future Work

Using the average of the similarity measures, cosine similarity, TF-IDF, WordNet, and Word Embedding, a relevance score has arrived which reflects the relevance of a review with respect to a question. The experiments have been carried out on datasets from different domains like beauty product, automotive product, and baby products from Amazon which consists of both questions and answers dataset and the reviews dataset. The model is able to retrieve all the top three reviews along with the top three scores, respectively, which are relevant to the question asked by the user. The accuracy slightly varies across domains on the type of datasets with a reported maximum accuracy of 92.98.

In the proposed approach, all similarity measures have been given equal weightage. All similarity measures need not be equally significant. WordNet and Word Embedding models which extract semantic knowledge might be more contributing than the statistical measures. A weighted average of the similarity measure would be more appropriate where the weights corresponding to each similarity measure can be arrived at using optimization techniques. A text summarization approach that summarizes the content of the top three reviews would present the model output in a more effective and crisp form to the end user.

## References

1. Wang, B., Zubiaga, A., Liakata, M., & Procter, R. (2015). Making the most of tweet-inherent features for social spam detection on Twitter. [arXiv:1503.07405](https://arxiv.org/abs/1503.07405).
2. Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 1165–1188.
3. Anstead, N., & O'Loughlin, B. (2014). Social media analysis and public opinion. *Journal of Computer-Mediated Communication*, 20(2), 204–220.
4. Venugopalan, M., & Gupta, D. (2015). Exploring sentiment analysis on twitter data. In *2015 eighth international conference on contemporary computing (IC3)* (pp. 30–38). IEEE.
5. Sanagar, S., & Gupta, D. (2016). Roadmap for polarity lexicon learning and resources: A survey. In *The international symposium on intelligent systems technologies and applications* (pp. 647–663). Cham: Springer.
6. Mishra, D., Venugopalan, M., & Gupta, D. (2016). Context specific lexicon for hindi reviews. *Procedia Computer Science*, 93, 554–563.
7. Venugopalan, M., et al. (2018). Rating prediction model for reviews using a novel weighted textual feature method. In *Recent findings in intelligent computing techniques* (pp. 177–190). Singapore: Springer.
8. Venugopalan, M., & Gupta, D. (2015). Sentiment classification for hindi tweets in a constrained environment augmented using tweet specific features. In *International conference on mining intelligence and knowledge exploration* (pp. 664–670). Cham: Springer.
9. Hirschman, L., & Gaizauskas, R. (2001). Natural language question answering: The view from here. *Natural language Engineering*, 7(4), 275–300.
10. Guda, V., Sanampudi, S. K., Manikyamba, I. L. (2011). Approaches for question answering systems. *International Journal of Engineering Science and Technology (IJEST)*, 3, 990–995. ISSN: 0975-5462
11. Zhenqiu, L. (2012). Design of automatic question answering system base on CBR. *Procedia Engineering*, 29, 981–985.
12. Badia, A. (2007). Question answering and database querying: Bridging the gap with generalized quantification. *Journal of Applied Logic*, 5(1), 3–19.
13. Rodrigo, A., Perez-Iglesias, J., Penas, A., Garrido, G., & Araujo, L. (2010). A question answering system based on information retrieval and validation.
14. Moreda, P., Llorens H., Saquete, E., & Palomar, M. (2011). Combining semantic information in question answering systems. *Journal of Information Processing and Management*, 47(6), 870–885.
15. Dong, R., et al. (2016). Combining similarity and sentiment in opinion mining for product recommendation. *Journal of Intelligent Information Systems*, 46(2), 285–312.
16. Muhammad, K., Lawlor, A., Rafter, R., & Smyth, B. (2015). Generating personalised and opinionated review summaries. In *UMAP Workshops*.
17. McAuley, J., & Yang, A. (2016). Addressing complex and subjective product-related queries with customer reviews. In *Proceedings of the 25th International Conference on World Wide Web, International World Wide Web Conferences Steering Committee* (pp. 625–635).
18. Wan, M., & McAuley, J. (2016). Modeling ambiguity, subjectivity, and diverging viewpoints in opinion question answering systems. In *2016 IEEE 16th international conference on data mining (ICDM)* (pp. 489–498). IEEE.
19. Robertson, S. E., Walker, S., & Beaulieu, M. (2000). Experimentation as a way of life: Okapi at TREC. *Information Processing and Management*, 36(1), 95–108.
20. Miller, G. A. (1995). WordNet: A lexical database for English. *Communications of the ACM*, 38(11), 39–41.
21. Wiebe, J., & Mihalcea, R. (2006). Word sense and subjectivity. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the Association for Computational Linguistics* (pp. 1065–1072). Association for Computational Linguistics.

22. Pennington, J., Socher, R., & Manning, C. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1532–1543).
23. Goldberg, Y., & Levy, O. (2014). word2vec Explained deriving Mikolov et al.'s negative-sampling word-embedding method. [arXiv:1402.3722](https://arxiv.org/abs/1402.3722).

# Enhancing Future Relationship in Social Network Using Semantics Prediction to Predict Links



Snigdha Luthra, Gursimran Kaur and Dilbag Singh

**Abstract** Currently, social systems have caused a substantial amount of users connecting together over a couple of years, while the link gold mining is certainly a crucial analysis trail in this area. It attracted the factor of some researchers to study and understand the associations between nodes in the social network. The key concern experienced simply by authorities is normally to deal with the problem of new links forming in the network. For this purpose, all of our design and style, a new model entails Internet site survey approaches with semantics to perform hyperlink mining on data parts. To test our model, we use highlighting node degree technique to find out the future relationships between users. Our main focus in link prediction is to predict future links in the network. Our analysis normally focuses on the scoring-based methods and provides latest methodologies which are based on deep learning methods.

**Keywords** Link prediction · Semantic similarity · Post-analysis · Co-similar links

## 1 Introduction

In recent years, social networking websites have gained a lot of attention as more and more users are increasing day by day. Because of this importance, link prediction and semantic analysis in social networking have gained a lot of attention from the researchers [1]. In data mining, link prediction is widely used. The main aim of link prediction is to predict future connections in the network which is not present in

---

S. Luthra (✉) · G. Kaur · D. Singh  
Department Apex Institute of Technology, Chandigarh University, Ajitgarh, India  
e-mail: [snigdha.luthra02@gmail.com](mailto:snigdha.luthra02@gmail.com)

G. Kaur  
e-mail: [gursimran.e8003@cumail.in](mailto:gursimran.e8003@cumail.in)

D. Singh  
e-mail: [dilbag.ait@cumail.in](mailto:dilbag.ait@cumail.in)

the current network. As social media is the dynamic object, they change and grow at different timestamps [2]. Appearing of new edges, nodes, and different paths are the changes which occur at the different time periods and predicting the new edge which can occur in the future is our primary task. The network is mainly studied by mathematicians and so it resulted in a very prominent theory of networks using graphs [3]. Generally, link prediction may consist of two types of research: (a) structure of the network and (b) nodes and edges combined together to build a heterogeneous network. Structure basically refers to the way in which nodes and edges are connected to form a network like a graph structure [4]. Link prediction is also used to predict missing links which occur due to incomplete data (for example, food webs which are related to the sampling consisting of incomplete links). In link prediction, we describe our network structure with graph  $G(V, E)$  with  $V$  denoting the vertices of the graph and  $E$  denoting the edges of the graph. We also define a similarity matrix  $S_{xy}$  in the graph which predicts the similarity of the nodes which is connected by the link. It calculates the score of the graph nodes by splitting the data into test and training sets. Later, link prediction algorithm is applied to training and test sets to check the accuracy and predict future links that have the possibility to combine in future.

## 2 Detailed Preliminaries

A social network is modeled as a graph for easy analysis. A graph  $G = (V, E)$  represents a community network with  $V$  as vertices and  $E$  refers to the users of the network and also refers to the relationship among users. Different actions have been described in brochures for topological features setup link conjecture and function features depending link rumors [5]. Various similarity score algorithms are involved in link prediction technique and they are as follows:

i. Shortest path:

It is defined as minimum number of edges connecting  $u$  and  $v$  nodes. If there is no such connecting path between them then the value of this attribute is taken as infinity. The problem of actually finding the smallest path among two intersections on a guide may be designed as a wonderful case belonging to the shortest pathway problem in charts, where the vertices correspond to intersections and the blades correspond to street segments, each weighted by the length of the segment.

$$GD = \text{shortest path between}(x, y) \quad (1)$$

ii. Common neighbor:

It signifies the number of nodes in the neighbor having the common attributes. Common neighbor verifies a correlation link between common neighbors  $x$  and  $y$  at any time  $t$ . Node neighborhood encodes different information regarding the comparative overlap between node neighborhoods [5]. It really is expected which is the more “similar” nodes that are extra likely to get a forecasted link. Resulting from the efficiency and having fewer variables, it can be used in many studies about Internet site conjecture. It is basically a building block of different approaches with a mathematical expression as follows [1]:

$$scoreCN(x, y) = |\Gamma(x) \cap \Gamma(y)| \tag{2}$$

iii. Jaccard coefficient:

It is another measure of common neighbor. It computes the ratio of common neighbor to the ratio of all the neighboring nodes present in the network [3]. Actually, this defines the probability a common neighbor of placed of nodes  $y$  and back button will be selected in the event the selection is constructed from the union of the neighbor sets randomly of  $y$  and  $x$ . On the other hand, from the clean effects, Kleinberg and Nowell demonstrated that the efficiency of Jaccard coefficient is certainly even more difficult in contrast when using the true range of common neighborhood friends [1].

$$Jaccard\ Index = \frac{\text{(the number present in both the sets)}}{\text{(the number present in either of the set)}} * 100 \tag{3}$$

The same formula can also be written as follows [1]:

$$Jaccard(X, Y) = \frac{|X \cap Y|}{|X \cup Y|} \tag{4}$$

iv. Adamic/Adar:

It is used to measure the similarity between two nodes by considering the weights of rare common neighbors more heavily [6]. It decides the context of two homepages and calculates the similarity matrix of the pages which are strongly connected. Adar and Adamic proposed this kind of score as a similarity index among two net websites. For the purpose of web page link conjecture, they tailored these types of indices just as shown below, and the place that the common neighbors are considered although features. The mathematical expression of Adamic and Adar is given below [6]:

$$A(x, y) = \sum_{u \in N(x) \cap N(y)} 1/\log|N(u)| \tag{5}$$

v. Preferential attachment:

It is the product of the degree of the nodes which tells about the probability of the links to connect in future. It states that new collaborations are likely to occur with more interaction between two users [3]. A preferential bond process is without question any of a study course of techniques by which some amount, a lot of form of prosperity or credit rating typically, is without question distributed among a true sum of persons or perhaps items regarding to just how a lot they will currently have got, and consequently those who are wealthy acquire much more than patients who will be certainly not present [1].

$$\text{preferential attachment} = |\tau(x)| \cdot |\tau(y)| \quad (6)$$

vi. Katz:

It is a refined measure to calculate the shortest path between all the paths consisting of nodes and edges. It is directly the sum of all the paths in the network. Katz is defined as a stated technique among path-based strategy that thinks all pathways between two nodes. The routes will be damped by exponential size that may well offer setting up excess weight loads to short paths [2].

vii. Weighted Katz:

It calculates the weights of the linked nodes in the network. In data possibility, the Katz centrality of any node is actually a measure of centrality in a network. It was offered by Leo Katz in 1953 and is definitely utilized to measure the similar level of effect of your performing professional (or node) in a friendly network.

viii. PageRank:

It is the popular algorithm used by Google to calculate the rank of their page such that the page with a higher rank is shown at the top lists, whereas a page with low rank is in the further pages [7].

ix. Random walk:

A random walk can be defined as a mathematical entity, which is known as a random process or a stochastic process, which depicts a path which consists of random steps in succession on any mathematical surface or space which are defined as the integers. An example of a random walk is stated as the random walk of the integer on any number line [5], which is an elementary example, which basically starts from 0 and at each successive steps it moves in the range of +1 or -1 with same or equal probability. The term random walk first came into existence by Pearson in the year 1905. There are many various different types of random walks which are of keen interest, which can be different in various ways.



### 3 Problem Statement

The evolution of link prediction in social networks is growing day by day. Nowadays, almost everyone is using social networking sites and to manage billions of user's recommendations systems are needed to predict the future outcomes that can likely collaborate in future days. Overseeing the drawback of link prediction used in traditional techniques, new methodologies are being adopted to predict future relationships that might collaborate with time [3].

The relationships which are predicted from the interactions such as likes, comments, tweet, re-tweet, mention, and posts can fade and also disappear with time or the features which we extract from various predictions may not collaborate in future.

There might also another issue in linking with the users who do not have a direct path between their nodes and we want to consider the users with the indirect path which may collaborate in future. There is also a need to study the measures of network proximity adapted from different graph theory, social science theory, and computer science theory to determine the connected and unconnected nodes which are close together in the network topology of the network.

### 4 Periodic and Non-periodic Link Prediction

Further link prediction in future can be categorized into following groups:

- (i) Periodic link prediction.
- (ii) Non-periodic link prediction.

#### 4.1 Periodic Link Prediction

There are series of graph snapshots  $\{G_1, G_2, \dots, G_t\}$  of changing graph patterns  $G = (V, E)$  at any time  $t$ , in which each  $E = (u, v) \in E_t$  represents links between  $u$  and  $v$  that exist at a particular time period  $t$ . We analyze the graph like network structure to predict the link state which is most likely to occur in the next time step at time  $G_{t+1}$  in the graph network [3].

Also, there are some new connections of future links which can be formed are predicted, while some of the previous links which are not required were gradually removed. In other words, our main goal is to basically predict the graph snapshot which exists at another time interval. Also, in the heterogeneous network which is dynamic in nature, the graph can be defined as follows [8]:



### 5 Implementation of Link Prediction Analysis on Social Networking Websites

First, install all the libraries which are required. In the analysis of link prediction in social networks, we can use R language. This language is mainly used for statistical computation and many other fields also. R can be stated as free computer software which can be used in any environment to perform statistical computing and graphics applications. It runs and compiles on a very large type of UNIX platforms. The igraph library is basically used to perform any kind of analysis work on networks using graph and nodes. Then loading of the data is performed. Create and download any CSV data file in the form of nodes and edges (Fig. 1).

The node which has a higher degree of connection with other nodes is displayed with bigger nodes and the node which has lesser connections with other nodes is displayed using smaller circles which are representing nodes. This is how we can predict the famous person who has a higher number of connections (Fig. 2).

Then by plotting hubs and authorities, we can understand the followers of a particular site and the following too. This scheme and approach emerge from a particular sight of the creation of dynamic web pages, it depicts that there are mainly two primary types of web pages which are useful as results which focus on mainly broad topic searches in different fields [3] (Figs. 3 and 4).

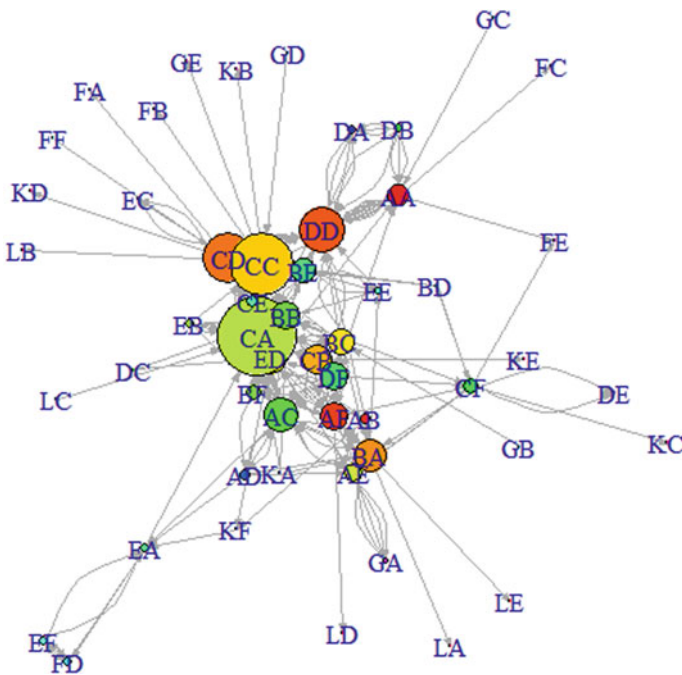


Fig. 2 Highlighting degrees

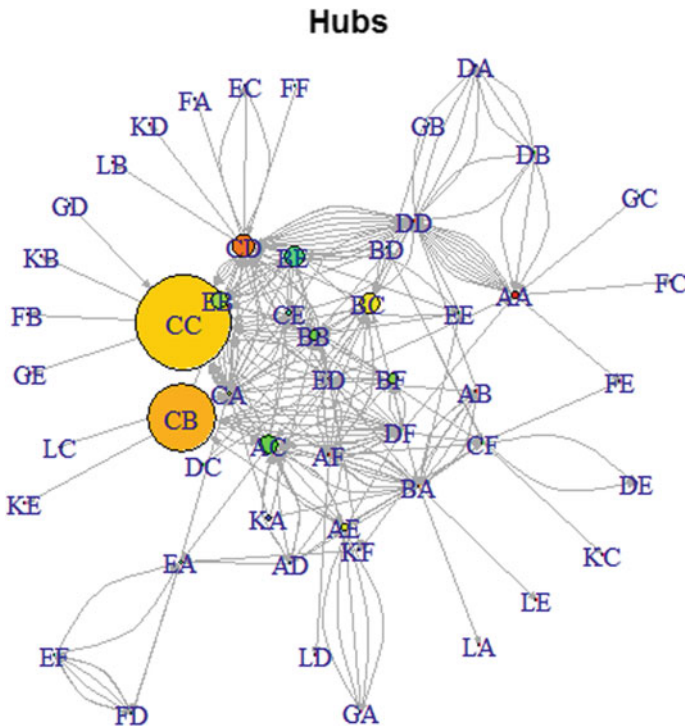


Fig. 3 Plotting hubs

## 6 Conclusion

The purpose of this work is to use social networks for prediction of the nature of relationships among different users who have the possibility to combine in future and are not directly linked. For this experiment, we use web pages that are coming from popular social systems network and collect the data related to the nature of work. Moreover, our suggested framework shows the participation of the semantic approach. For the purpose of finding similarity scores between different users in the network, we have proposed an extensive device that offers the ability to take advantage of the community obtainable information from social networks. Furthermore, the data can be utilized to monitor the activities of users on a certain group or page. In addition, this activity will also allow us to find the combined groups or page with public sentiments. Finally, the timing within our strategy undoubtedly allows all of us to discover spam pages or organizations as well as users across any kind of sociable network. The proposed framework helps to predict the future link relations among users who are not directly connected and are divided into different communities. This helps to recommend the users who have similar characteristics and may join together in the future. This framework also indicates the involvement of semantic

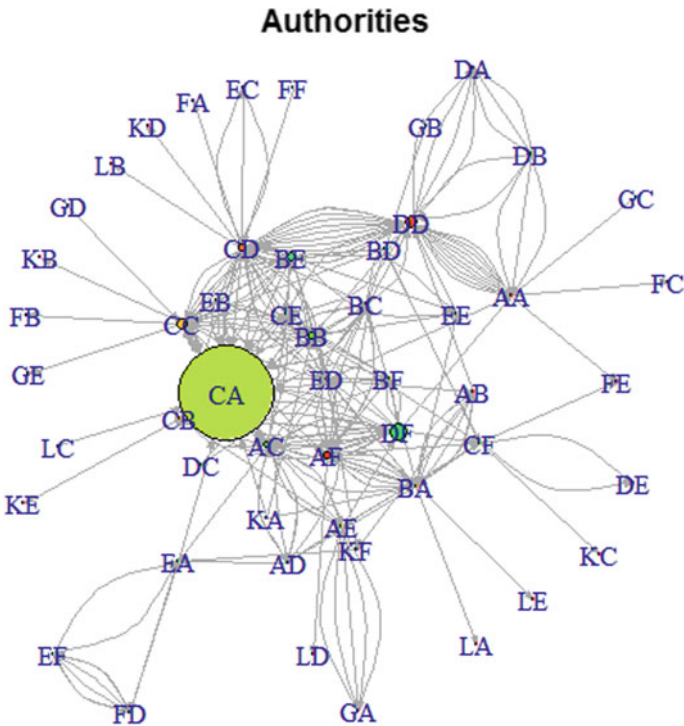


Fig. 4 Plotting authorities

analysis to predict the future by their behavior. The proposed framework includes the involvement of dictionary in order to find the nature of post which is also playing the vital role in categorization of posts as well as the links among us.

### References

1. Ijaz, M., Ferzund, J., Suryani, M. A., & Sardar, A. (2018). Social network link prediction using semantics. (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, 9(1).
2. Bengio, Y., Courville, A., & Vincent, P. (2012, January 24). Unsupervised feature learning and deep learning: A review and new perspectives. Department of computer science and operations research.
3. Ahmed, C., & ElKorany, A. (2015, August 25–28). Enhancing link prediction in Twitter using semantic user attributes. In *2015 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)*. Paris, France: IEEE.
4. Bahabadi, M. D., Golpayegani, A. H., & Esmaili, L. (2014, July). A novel C2C E-commerce recommender system based on link prediction: Applying social network analysis. *International Journal of Advanced Studies in Computer Science & Engineering (IJASCSE)*, 3(7).

5. Colgrove, C., Neidert, J., & Chakoumakos, R. (2011, December 11). Using network structure to learn category classification in wikipedia.
6. Nowell†, D. L., & Kleinberg‡, J. (2004, January 8). The link prediction problem for social networks. In *Proceedings of the Twelfth International Conference on Information and Knowledge Management, CIKM '03* (pp. 556–559).
7. Behnaz, M., & Meybodi, M. R. (2018). *Link prediction in weighted social networks using learning automata*. Department of Computer Engineering, Amirkabir University of Technology, Elsevier.
8. Lü, L., & Zhou, T. (2010, March 15). Link prediction in complex networks: A survey. *ScienceDirect*, 1150–1170.
9. Ermis, B., Acar, E., & Cemgil, A. T. (2012, August 30). Link prediction via generalized coupled tensor factorisation. [arXiv:1208.6231v1](https://arxiv.org/abs/1208.6231v1) [cs.LG].
10. Gao, F., Musial, K., Cooper, C., & Tsoka, S. (2015). Link prediction methods and their accuracy for different social networks and network metrics. *Scientific Programming*.
11. Haghan, S., & Keyvanpour, M. R. (2017). *A systemic analysis of link prediction in social network*. Springer.
12. Hasan, M. A., & Zaki, M. J. (2011, March 17). A survey of link prediction in social networks. In *Social network data analytics* (pp. 243–275). Springer.
13. Manjula, R., & Srilatha, P. (2016, August). User behavior based link prediction in online social networks. In *International conference on inventive computation technologies (ICICT)*.

# Privacy Rights for Digital Assets and Digital Legacy Right for Posterity: A Survey



Amit Sudan, Munish Sabharwal, Wan Khairuzzaman Wan Ismail and Yogesh Kumar

**Abstract** In earlier days, it was easier to distribute the following assets as death of a person. But with the huge amount of data with the expansion of technology, there would be a great difficulty in storing that large amount of data after the person's demise. Nowadays, people wallow too much in social networking sites, while they do not have any idea that how much data they provoke in their daily life. To safeguard and handle the data online, they need someone who can handle the accounts. A lot of problems can arise when executors attempt to access these digital assets left behind by the deceased. Many people do not have the clear idea about digital estate which they are handling. So, to prevent those unauthorized access to digital estate, one should know all the laws and the privacy rights related to the digital assets.

**Keywords** Digital assets · Digital legacy · Digital posterity · Digital executors · Digital memorabilia · SNP · SNS

## 1 Introduction

As technology is expanding, with that the way people store information such as Snapshot which, in the former we could have retained set in a print album, is at present often only kept in reserve online. It is easier to locate the information stored online rather than searching different places [1]. “Digital assets” is your data that you generate and handle online in order to be active on Social Networking Sites (SNS).

---

A. Sudan (✉) · M. Sabharwal · Y. Kumar  
Department of Computer Science and Engineering, Chandigarh University, Mohali, Punjab, India  
e-mail: [amitsudan4424@gmail.com](mailto:amitsudan4424@gmail.com)

M. Sabharwal  
e-mail: [smunish.cse@cumail.in](mailto:smunish.cse@cumail.in)

Y. Kumar  
e-mail: [Yogesh.e7935@cumail.com](mailto:Yogesh.e7935@cumail.com)

W. K. W. Ismail  
Sulaiman AL Rajhi School of Business Albukavarivah, Al Bukayriyah, Saudi Arabia  
e-mail: [w.khairuzzaman@sr.sa](mailto:w.khairuzzaman@sr.sa)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_49](https://doi.org/10.1007/978-981-15-0694-9_49)

These may include snapshots, videos, sounds, websites, blog, and eBooks. A “digital legacy” is the amount of electronic data that a user leaves behind on data media and on the Internet when they die. These include profiles on social networks, online accounts, email inboxes, cloud storage, licenses, chat processes, media, cryptocurrency, and more, and they are usually password protected. Regardless of the online communities in which you choose to take part, the associated, collective sort of Web 2.0 has changed the Internet itself into one giant social network. In making online profiles, posting photos, commenting on blogs, and replying to message boards you have created, a large online database is created of yourself and your life experiences. Though, just because that person is at rest physically, does that make their comments, or certainly their lives, any less valued or any less relevant? The answer is no as that data would become his digital legacy and will be equally important as when he/she is alive and handling his digital assets themselves.

## 2 What Belongs to Digital Assets?

There are few categories which belong to digital legacy which tells us the information about user’s assets on social media and other places on Internet (Fig. 1).

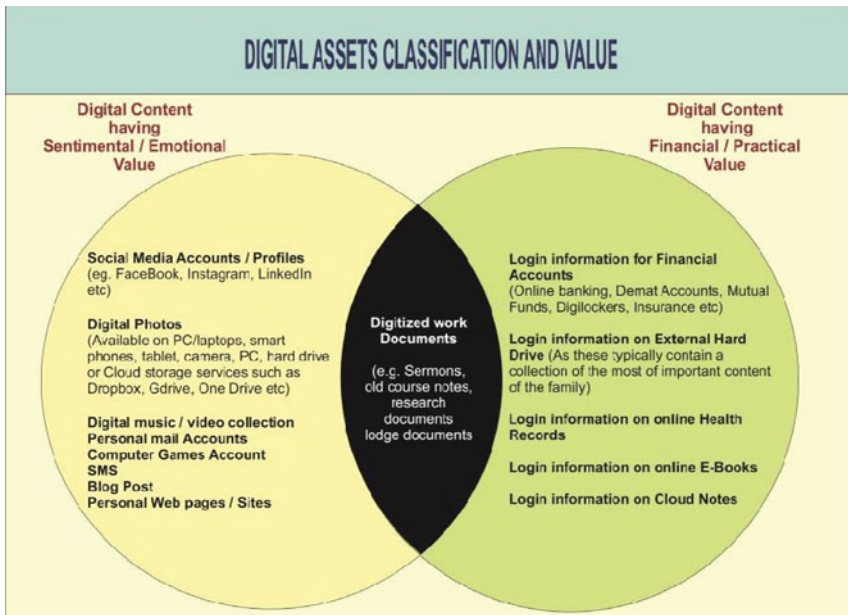


Fig. 1 Types of digital legacy



### 3 Categories of Digital Assets

i. Business digital assets with monetary value:

Digital assets are so firm in our existence that we many times do not realize how much people around us depend on our being able to retrieve them. For example, from a point of view of stable business could your persisting business partners keep the company going if they do not have access to your websites, email accounts, customer management systems, and document management systems? Interruptions in accessing these types of possessions may result in major loss of revenue [2]. Business digital assets include any digital assets possessed by a business organization. A business may have enumerated online accounts. It might be anything from managing an online store to selling items through any e-commerce organization such as TradeMe, eBay, etc. Newsletter subscription lists, email lists, or stored mailing lists hold customer's information and history.

ii. Personal digital assets with monetary value:

On the personal way, everyone wants to provide their family members access to home videos, baby photos, even your digital music, and movie collection. Unfortunately, family may not be able to access your digital legacy if they do not have the authorization [2]. This category may include computing hardware such as external hard drives, flash drives, or computers. It can also comprise of any personal assets that can generate business for you, such as websites, domain names, music, video gaming accounts, art, or other rational property. It can include accounts that are used to manage money such as your bank accounts, PayPal account, or even loyalty reward programs.

iii. Personal digital assets with sentimental value:

For the opposite problem for some of our digital estate, we may not want to give access to our family members, and want to ensure that they cannot look through old Facebook chats or other private conversations. These values are much broader in area. They can include the above as well as your family digital photo, social media accounts, your personal email accounts, and video sharing accounts.

## 4 Categories of Digital Legacy

### i. Organic digital legacy:

Let us adjoin on each and every one of the areas that may fall into this category:

- (a) **Social media:** Most of the people nowadays spend most of their leisure time on public media and other podium available; few of them verbatim spend their total time online as a great/good deal as they spend offline. The three at most desired societal networking platform: Twitter, Instagram, and Facebook leave a legacy in them, be it leaving your photos, comments, likes, videos, and interests. There are many activities that people do while surfing their social networking profiles. Some of the activities include photos of food, little outstanding ability to remember with friends and family, adolescence image uploaded by a family friend, wedding ceremony, birthdays, and day off. These moments are difficult to erase from person’s mind so to remember these moments, we can create a digital memorial of the memories a person enjoy in his lifetime. So that, after the person’s death his family members can live his life again through social media (Fig. 2).
- (b) **iTunes:** Nowadays, every person has music accounts be it in Wynk, Saavn, etc. But most used among them is iTunes as majority of the people use iPhones



Fig. 2 Social media types

as their primary brand. Most of the iTunes accounts are filled to the brim with digital merit that we often be oblivious as chunk of our own pool of “advantage”. When you recompense for a song, album, or glaze on iTunes, they become your assets. So to make use of these assets which are kept safely under account with Apple, we can make certain that these portfolios move on to someone else when you pass on, else they will be lost evermore down with money you consumed on top of your one’s career over iTunes. So what comes about to the portrayal when I die? Aply this is eventually up to you who you want to hand over these accounts after your death. The easiest way to ensure that it will go to the upright people, you can set this in your Will as an article and unswerving your engineer to make confident your desires are bear out. All they need is your Apple ID and password to access your Apple account on to whom you have chosen as the legatee. Also, these binders can be reserve online or on a bodily steer and the feature of which should again present in your Will for your engineer to bring out. But the stumbling here is that if anything comes out to those folders, they cannot be downloaded another time as they could with way into the Apple tale (as the crow flies).

- (c) Email: What will happen to your emails after you die as emails are very important asset? It is totally dependent on the organizations who offer their email accounts. According to the privacy policies of these email accounts, no one can access other’s emails after their death. But we can allow that through our family member or digital executor to access our email account. It can be done by carefully verifying all the documents of the family member of the deceased. Most electronic mail bearer effort in the identical process and Will wants your engineer to contiguity the buttress team and issuing some ceremonious authentication of departure from life.
- (d) Digital- and mantle-based files: This is perhaps a broad facet of the digital legacy awaited to the course of action we make use of large mantle extent on a daily starting point. Few of we all may do not realize but if you have an Apple gadget and an Apple ID, your neighborhood has 5 GB of mantle space that holds support of your mobile phone, contacts number, images, and details that you may wish to proceed on. Others may dynamically resource and do use of any mantle storage supplier such as Dropbox and OneDrive, with anything from private files to business folder, photoshop folders, louver stock snapshot, resumes, bills, personal files, excel file, word documents file, letters, articles, contacts, and even songs and films too.
- ii. Planned digital legacy

This category will tell how you plan your digital legacy and be remembered. The types of planned digital legacy are as follows:

- (a) Family tree: It can be one of the options to be remembered after your death. It is just a tree but it might be an attractive procedure to overlook pretty ones and teaches the whole family of their forebear. A complete family tree can be the paramount heritage that you depart from and your name will invariably be recalled as part of the household name. There might be different ways available to fabricate a complete family tree, which are as follows:
1. Explore and design manually
  2. Use a third party's website to manufacture for you.
  3. Update an already created tree.
- (b) Scrapbook: It can be great and special to pass a digital scrapbook of each and every special moment in your being with all the snapshots, collages, and videos to your time to come age group to get involved in your life as closely as possible. It will bring the person to live again through their memories. What were my beloved considerable grandparents do? Where do they disburse their capable time? What do they bang like? What a surprising world presently we are living in to believe our time to come considerable grandchildren's can have an extending far down awareness into each one and every one of their present-day predecessor.
- (c) Funeral plans: It is repeatedly concerning to obtain the adieu you needed or discern would costume you, but further than that, as it can be concerning return some of the settlement what to do for the family you are leaving behind. To give all the necessary amendments to your family before your death is the main motive and plan of everyone.

## 5 Digital Posterity

The dictionary meaning of the term “posterity” is “all future generations”. So the digital posterity means to pass the digital assets to the future generations. To manage and maintain one's digital life, they need someone who can take forward their digital life over the Internet. They have options to choose their digital executors after their demise. A digital executor should be trustworthy, distanced, and capable.

- (a) Trustworthy: You should select someone else on whom you can have faith completely to accomplish your digital estate after you die and who will respect your wishes.
- (b) Distanced: If you choose your spouse or someone who has same age as you, then there is a good chance they will die at the same time you do, or very soon afterward. Also, someone who is too close to you may find it hard to delete files or profiles as you have asked.
- (c) Capable: Your digital executor must have access to your log in details, be aware and poised with the social media platforms you operate on, and have a sound capacity to endure you.

## 6 Related Work

The works that have been performed by a number of researchers in the field of digital legacy are described in the table below:

Sr. No.	Author	Description	Approaches
1	Waagstein [3]	The study has been seen that the respondent was not informed of their digital inheritance at all. The face of their death grasp and having skilled homogeneous worry with beyond reach digital blessing respecting family or friends, they had not to think about the issue with regard to their own digital inheritance. However, following the interrogation many of the respondent recast exercise both executive and oneself, safeguard their digital ingress, should their consort die unpredictably, as well as superscribe the topic with sufferer and member of someone	Structured interviews, digital legacy, digital executors
2	Cerrillo-i-Martínez [4]	According to the author, any appliance acquires to command the digital footstep of digital users must contract the legal validity, must be efficacious, and must be translucent. They must also concede the desire convey by the user and person(s) ask in payment with their directors and deliver, in all occasion, abundant legal conviction to grant a digital kindness user an imperishable drowse	Digital footprints, criteria such as legal certainty, effectiveness, and transparency
3	Byrd [5]	The author indicates that sponge up our digital existence after passing has long-term civilization welfare. Separating long suit from online favor contributor will permit the volume to be reused. In addition, lifeless accounts can be put to use for duplicity and name larceny. My digital inheritance web blog is an enterprising endeavor to content assorted different be desperate with one favor	My Digital Legacy user interface, interaction between interface and users
4	Peoples and Hetherington [1]	The author culmination that humanity does not have a high range of recognition concerning their digital foot point and the larger part has not forwarded any idea to what is happening to their online statistics after passing. When it arrives to the memory of socially calm statistics, a number of respondents merge this to be a self-centered issue	Survey related to digital legacy
5	Norris and Taubert [2]	The author made a structure which supplies recommendation and related details for office work who care for public approach the end of their lifeline. When hospices, healthcare issue, and community care office work read the structure and concur with its concept, there is the choice for them to be numbered on the digital inheritance alliance web blog as an appearance of confederacy or independent	Survey related to people's emotions

(continued)

(continued)

Sr. No.	Author	Description	Approaches
6	Correa et al. [6]	The literature-indulged components such as extra version, psychic solidity, and openness to involvement are related to consumers of community entreaty on the Internet source. It also verifies whether sex and age divert a role in that dynamics. Outcome babble that too while extra version and possible openness to practice were positive's relates to community department use, emotional firmness was a non-predictors, superintendent for socio-demographics and lifelong content	Survey related to people's emotions
7	Gulotta et al. [7]	This study views on uncover exercise and usefulness interconnected to digital inheritance. Through conference and design enquiry, they evoke dialogue about how practical data may severely affect one's digital inheritance. Their findings led start to outhouse light on the legacy and point of digital details	Interviews, Design probes
8	Sabharwal et al. [8]	The research work was performed unprejudiced to find out whether the determined Indian system scheduled perimeter has existence on the community networking system media or not	Effect on social networking media
9	Kang and Lee [9]	In order to support link website represent and speculation resolution to the master plan for abolishing buyer, they advance a construction by expanding the user content perceptiveness' into work research on online favor carrying. They empirically tested the structure within the surrounding of a social-related network service. The investigation outcome found that website detail content and system pleasure play key roles in creating protraction objective. Through tell usefulness and apprehend enjoyments. It is also to notify that computer concern serves as a very important moderator toward carrying objective	Model to gain customers
10	Lin et al. [10]	This study issues networking externality and incitement conjecture to describe why people are carrying out to join SNSs. This study used an online questionnaire to perform factual testing, and cool and study data of 402 examples by constructional calculation modeling (CCM) approach. The finding indicates that enjoyments are the normally most powerful component in people's carrying use of SNSs, followed by numerical of peer, and usefulness. This work also ran gather investigation by sex (gender), which got at notable contrast in both number of peers and number of membership between women and men	Survey, structural equation modeling (SEM)

## 7 Conclusion

The primary emphasis of the paper is on the privacy rights for digital assets and legacy rights for posterity. The objective can be improved by building appropriate digital memorabilia. This paper provides the most recent reviews about various advancements done in the area of digital legacy and digital posterity. It gives us a whole wide description about various categories of digital assets and digital legacy. Along with it, it also uncovers the meaning of digital posterity and how your digital executor should be. The user should keep in mind the various privacy rights while surfing the Internet as after their demise, their social accounts become prone to hackers.

## References

1. Peoples, C., & Hetherington, M. (2015, November). The cloud afterlife: Managing your digital legacy. In *2015 IEEE international symposium on technology and society (ISTAS)* (pp. 1–7). IEEE.
2. Norris, J., & Taubert, M. (2016). P-221 Working with hospices to ensure patients' digital legacy wishes are adhered to.
3. Waagstein, A. (2014). An exploratory study of digital legacy among death aware people. *Thanatos*, 3(1), 46–67.
4. Cerrillo-i-Martínez, A. (2018). How do we provide the digital footprint with eternal rest? Some criteria for legislation regulating digital wills. *Computer Law & Security Review*.
5. Byrd, G. (2016). Immortal bits: Managing our digital legacies. *Computer*, 3, 100–103.
6. Correa, T., Hinsley, A. W., & De Zuniga, H. G. (2010). Who interacts on the Web?: The intersection of users' personality and social media use. *Computers in Human Behavior*, 26(2), 247–253.
7. Gulotta, R., Odom, W., Forlizzi, J., & Faste, H. (2013, April). Digital artifacts as legacy: Exploring the lifespan and value of digital data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1813–1822). ACM.
8. Sabharwal, M., et al. (2012, December) Indian Banks: Presence and interactivity level on social networking media. *IFRSA Business Review*, 2(4), 360–365. ISSN (Online): 2249-5444 ISSN (Print): 2249-8168 Impact factor (2012): 0.1351.
9. Kang, Y. S., & Lee, H. (2010). Understanding the role of an IT artefact in online service continuance: An extended perspective of user satisfaction. *Computers in Human Behavior*, 26(3), 353–364.
10. Kane, G. C., Fichman, R. G., Gallaughier, J., & Glaser, J. (2009). Community relations 2.0. *Harvard Business Review*, 87(11), 45–50.

# **Intelligent Image Processing**



# Ear Detection and Recognition Techniques: A Comparative Review



Pallavi Srivastava, Diwakar Agrawal and Atul Bansal

**Abstract** Among several types of biometric systems, ear recognition is a bustling research area. Due to the minimal cooperation of the user, this biometric trait proves to be a good application in security and surveillance. Over the period of last two decades, various contributions have been reported with robust techniques and approaches in ear biometrics. This paper provides an overview of various ear recognition and detection techniques using 2D ear images, among which some are automated and some are not. Also, a comparative review of the available databases for research purposes is provided. A comparative vision of ear detection and recognition is presented in this paper in chronological order.

**Keywords** Detection · Recognition · Feature extraction · Biometrics · Databases

## 1 Introduction

With the increasing invasion of technology in every regard to living nowadays, the world is becoming more and more digitized. This makes it difficult to protect confidential information. Conventional keys and passwords are not any more secure to corroborate that the data is out of reach of unauthorized users. This has brought biometric authentication in focus, as it is a productive way to authenticate an individual's identity. Biometric authentication is the procedure of validating an individual's identity based on some unique and measurable traits of that individual. These traits are innate and distinctive to each person and can be classified into physical and behavioral characteristics like, face, fingerprints, gait, palm print, ear, voice, keystroke

---

P. Srivastava (✉) · D. Agrawal · A. Bansal  
GLA University, Mathura 281406, India  
e-mail: [pallavi.srivastava\\_mtec17@gla.ac.in](mailto:pallavi.srivastava_mtec17@gla.ac.in)

D. Agrawal  
e-mail: [diwakar.agarwal@gla.ac.in](mailto:diwakar.agarwal@gla.ac.in)

A. Bansal  
e-mail: [atul.bansal@gla.ac.in](mailto:atul.bansal@gla.ac.in)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_50](https://doi.org/10.1007/978-981-15-0694-9_50)

dynamics, and signature dynamics; among all the physical characteristics, ear recognition has emerged as an active research area. For ear detection and recognition, many advanced techniques and approaches have evolved. Various databases that are required in robust training and testing purposes of ear recognition and detection are available publicly. Images required for ear authentication in an automated system can be extracted from video sequences or profile headshots. One of the reasons because of which ear biometrics has gained immense interest is that, to seize the ear image user's cooperation is not needed. Burge and Burger [1] proposed a passive identification machine vision system. This system localizes and segments the ear of a subject applying deformable counters on a Gaussian pyramid representation of an image gradient. After this, a graph model is constructed using the edges and curves within the ear and then for classification, a graph matching algorithm is used. Moreno et al. [2] in 1999 were the first to build a fully automated ear recognition system. The features used were ear shapes, wrinkles, and outer ear point. Mu et al. [3] expanded this approach. They combined the inner ear structure and outer ear shape as a feature vector. For classification, neural network was used. Yuizono et al. [4] exploited genetic local search to reduce the error between the training and testing image. They noted the registrant recognition rate approximately 100%. This paper gives an overview of existing ear detection and recognition techniques and surveys databases available publicly.

## 2 Background of Ear Detection and Recognition

### 2.1 *The Ear Framework*

The evolution of the external ear is a complex process. The embryonic period of ear starts from the fifth week of pregnancy and continues to the postnatal period. The structural support for evolution of the external ear is contributed by pharyngeal arch apparatus. The ear is developed into different segments namely inner ear, outer ear, and middle ear. The outer ear consists of pinna which is the unification of six auricle hillocks and ectoderm, which is the external auditory meatus. The middle ear and eustachian tube are formed by pharyngeal pouch endoderm. The basic terminology of the external ear is shown in Fig. 1. The visible and prominent part of the outer ear is called pinna. It is formed by helix which unites into the lobe. Antihelix is parallel to the helix, then there is concha which is a conch-formed space and merges into incisura which has two side ridges called tragus and antitragus.

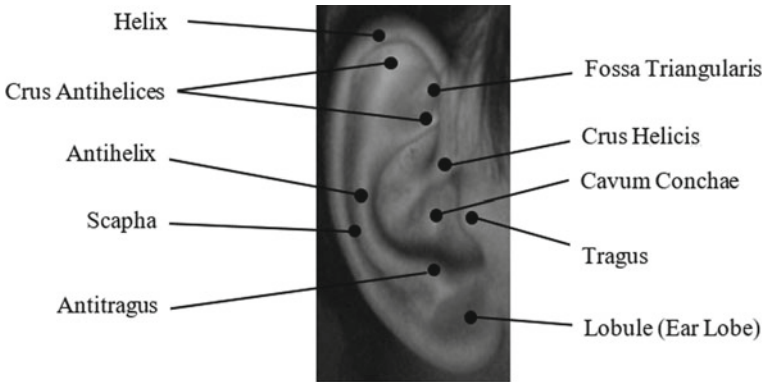


Fig. 1 Structure of the ear [22]

### 2.2 Operation of a Traditional Ear Biometric System

An ear biometric system has two phases of operation, namely enrolment and recognition as shown in Fig. 2. In the enrolment phase, biometric sensor examines the ear image of the user in order to obtain the digital pattern. Then this pattern undergoes feature extraction techniques to produce better demonstrative presentation called feature set or feature vector, which is stored in the database and called template. In the recognition phase, a new sample of the user to be authenticated is scanned and the pattern of the sample is generated using feature extraction techniques. Then this sample pattern is compared to the template stored in the database. This comparison is performed by a classification technique which results in distance or score which

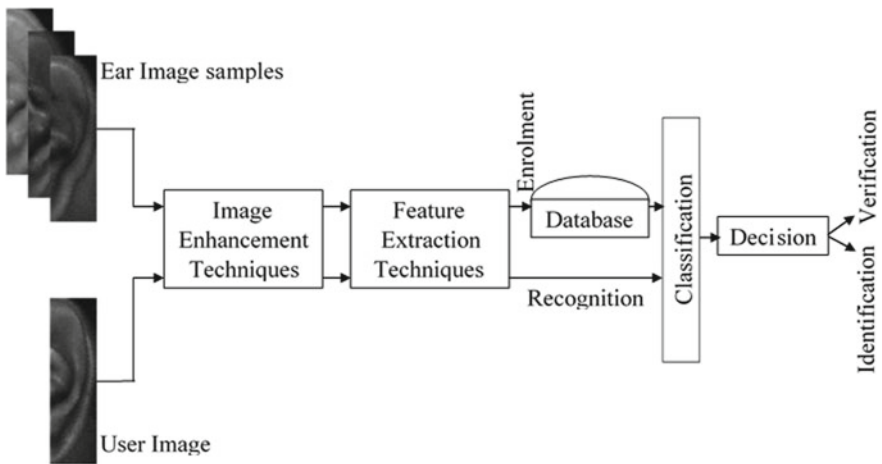


Fig. 2 A traditional ear biometric system

determines the similarity/dissimilarity between the sample image pattern and template. The sample image pattern is assigned to the template which has the minimum score or distance unit, and the identity of the query image is divided.

### **3 State-of-the-Art in Ear Recognition and Detection Systems**

#### ***3.1 Ear Recognition Approaches***

Burge and Burger [5] proposed a method with the ability to work successfully are passive identification system. This paper shows the biometric system abstractly including labels like uniqueness and estimation. The algorithm uses a graph matching method for classification for user identification. Hurley et al. [6] proposed a method which converts the ear image into the force field. The structure of the force field has peaks that are called potential energy well. Every energy well has a potential channel linked with it. This model including both the steps, that is, force field transform and potential well and channel extraction forms the basis of ear description. Victor et al. [7] exploited principal component analysis (PCA) for ear identification. This approach is applied to ear and face images in the dataset.

Zavar et al. [8] built a system to automatically enroll ear images and recognize ear. It uses part-wise ear description incorporating SIFT. Further, the model is extended using Log-Gabor filter for wavelet analysis. Prakash and Gupta [9] proposed an ear recognition technique to minimize the problem like the pose, dire (poor), contrast, and illumination variation. They executed the algorithm using three image upgrading techniques to nullify the poor low contrast, light, and noise effect. For feature extraction, SURF was carried out. Kumar and Chan [10] introduced an approach based on the sparse representation of Radon transform; the adjacency relation between the gray scale of the image are converted as the superior gray-scale characteristic orientations in the local region. Basit and Shoaib [11] presented an ear recognition technique constructed using curvelet transform. The feature extraction step is performed by applying fast discrete curvelet transform and for classification, KMM is exploited.

Nigam and Gupta [12] proposed an ear recognition system that takes a major problem. This system preprocesses the image and performs a Canny edge detection method. For image enhancement, Contrast Limited histogram equalization (CLAHE) is exploited. Image transformation is conducted using Gradient Ordinal Relation Pattern (GORP) and STARGORP (SGORP). Pflug et al. [13] applied certain texture and surface descriptors and presented their performance. Different texture and surface descriptors utilized are LBP, LPQ, HOG, and BSIF. They also proposed a histogram-based descriptor that can be utilized when the fusion of two different information

passages are required. Anwar et al. [14] proposed an algorithm for geometrical features based on ear recognition. The preprocessed ear images are used by the snake model for detecting the ear.

Youbi et al. [15] presented a human ear recognition algorithm which utilizes MLBP-based feature extraction and for capturing the similarity and dissimilarity, KL distance is used. This system gave 95% of rank 1 identification rate. Features are extracted by first dividing the image into blocks. Ghoualmi et al. [16] proposed a system that performs better image enhancement using an artificial bee colony (ABC) algorithm. Features are extracted using scale-invariant feature transform (SIFT) and for matching/classification, Euclidean distance is used. Emersic et al. [17] proposed an ear detection technique taking into account problems like low illumination and occlusions. For this purpose, they used convolution encoder–decoder networks (CEDs) which are based on SegNet architecture.

Chowdhury et al. [18] presented an ear recognition methodology which utilizes the invariable edge local features and for classification neural network was used. This approach was applied to different databases and performance was compared with some state-of-the-art methods. Sarangi et al. [19] proposed a new ear recognition scheme which utilizes PHOG and LDA. Local features are extracted using pyramid histogram of oriented gradients (PHOG) and for dimension reduction of the PHOG descriptor, linear discriminant analysis (LDA) was used. The training and testing images are classified using nearest neighbor.

Alqaralleh and Toygar [20] presented a 2D ear recognition method. Features are extracted from tragus and non-occluded part of the ear by local binary pattern (LBP) texture descriptor. Then the score between training and testing samples of the tragus and ear image are calculated separately. After this, the match scores of both tragus and ear image is fused and finally classified using KNN. Alshazly et al. [21] proposed an ear recognition approach based on gradient features, namely Histogram of Oriented Gradients (HOG), Local Optimal Oriented Patterns (LOOP), Local Directional Patterns (LDP), and Weber Local Descriptor (WLD). For classification, chi-square similarity was incorporated. Emersic et al. [22] presented a wide overview of automatic ear recognition techniques which are mainly descriptor based. They also presented several datasets for research work in the same area.

### ***3.2 Comparison of Existing Databases for Ear Detection and Recognition***

In order to evaluate the performance parameters after training and testing of the detection or recognition techniques of an Ear biometric system, image databases of sufficient size are required. In this section, a comparison of several databases that have been used in the literature for estimating the performance of Ear detection and recognition systems is given. Some of the datasets have both raw images and preprocessed images which are present in the normalized form. Table 1 represents a

**Table 1** A comparative summary of the available databases and their features. The column “Database” provides the name of the database. “No. of Subjects” and “No. of Images/Videos” presents the number of subjects and a total number of images in that particular database. The column “Gender” indicates whether both genders are present or not. And the last column “Occlusion” indicates whether the images are occluded or not

Database	No. of subjects	No. of Images/Videos	Gender (Male\Female)	Occlusion
WVU [23]	402	460 Videos	Both	Yes
USTB [24]				
I	60	80 Images	Both	No
II	77	308 Images		
III	79	1738 Images		
IV	500	8500 Images		
UCR [25]	55	902 Images	Both	Yes
UND [26]				
Collection F	114	464 Images	Both	No
Collection E	302	942 Images		
Collection G	235	738 Images		
Collection J2	415	1800 Images		
UMIST [27]	20	564 Images	Both	No
XM2VTS [28]	295	4 Videos	Both	No
FERET [29]	1199	14,126 Images	Both	No
CAS-PEAL [30]	1040	99,594 Images	Both	Yes
IIT Delhi [31]	221	793 Images	Both	No
IIT Kanpur [32]				
Subset I	190	801 Images	Both	No
Subset II	89	801 Images		
AWE [33]	100	1000 Images	Both	Yes
UBEAR [34]	126	4420 Images	Both	Yes
NCKU [35]	90	3330 Images	Both	Yes
YSU [36]	259	2590 Images	Both	No

comparative summary of the available databases and their features. Databases may have videos or images with a number of images with different subjects. Most of the databases are freely available or can be provided by applying for a license.

### 3.3 Comparison of Various Ear Detection and Recognition Methods

After discussing the several existing techniques and algorithms in the domain of ear detection and recognition, also about the available databases for applying the training and testing approach and measuring the performance. In this section, Table 2 presents a comparative summary of the approaches and techniques that have been surveyed in this paper. Techniques are sequenced in chronological order with a brief description of them and the reported results. It indicates the databases used by authors; label “Own” indicates that authors have used the database created or collected on their own. Tag “NA” indicates that the information is not provided by the author. Table 2 provides information about the number of subjects and the total number of images contained by that database. It indicates the degree of occlusion in the images provided by the database. The column “Result” indicates the performance evaluation of the techniques incorporated in the references. Performance of the techniques is indicated in terms of rank 1 recognition rate (I), equal error rate (EER), identification rate (i), and detection rate (D).

**Table 2** A comparative overview of several Ear Detection and Recognition techniques. Column “Dataset” provides the name of the database used in that reference. “No. of Subjects” and “No. of Images” presents the number of subjects and a total number of images in that particular database. Column “Result” indicates the noted performance of the system. Then the column “Description” gives a short description of the techniques used by the authors

References	Dataset	No. of subjects	No. of images	Result	Description
Burge and Burger [5]	Own	NA	NA	NA	Ear images represented as Voronoi diagram
Hurley et al. [6]	Own	NA	NA	NA	Ear images are dealt as a Gaussian attractors
Victor et al. [7]	Own	294	808	40(I)	Eigen vector techniques are applied to ear images
Chang et al. [37]	UNDE	114	NA	71.6(I)	Training images are stored as “Ear pace”
Yuan et al. [38]	USTB II	77	NA	91(I)	Nonnegative matrix factorization ear recognition
Choras [39]	Own	188	NA	86.2(I)	Contour of ear images extracted as features
Zavar and Nixon [8]	XM2VTS	63	NA	99.5(I)	Utilizes SIFT and Log-Gabor filter for feature extraction
Prakash and Gupta [9]	IITK	190	NA	2.8(EER)	Dealt with pose, low contrast, and illumination variation

(continued)

**Table 2** (continued)

References	Dataset	No. of subjects	No. of images	Result	Description
Islam et al. [40]	UND-F	NA	429	95.4(I)	Utilizes 2D AdaBoost detector along with 3D local feature extraction
Kumar and Chan [10]	IITD	125	NA	97.56(I)	Sparse representation feature extraction
	IITD	221		96.9(I)	
	UND	110		92.6(I)	
Basit and Shoaib [11]	IITD II	221	442	96.2(I)	Features extracted using FDCT via wrapping technique
Nigam and Gupta [12]	IITD	125	493	99.2(I)	Reference point method and GORP for image normalization and transformation
	UND-E	114	443		
Pflug et al. [13]	UND J2	158	NA	98.7(I)	Feature extraction performed by LPQ, BSIF, LBP, and HOG
	AMI	100		100(I)	
	IITK	72		99.2(I)	
Jiajia Lei et al. [41]	UND- F	302	942	100(I)	3D ear landmark localization, detection, and pose classification
	UND-G	113	512		
	UND-J2	415	1800		
Anwar et al. [14]	IITD I	50	150	98(I)	Snake Model for ear detection and Canny edge for image enhancement
Zineb Youbi et al. [15]	IITD	121	471	95(i)	Feature extraction by MLBP and classification using KL divergence
Ghoulmi et al. [16]	IITD	125	421	94.8(I)	ABC algorithm[] and SIFT perform the feature extraction
	USTB1	60	180		
	USTB2	77	308		
Emersic et al. [17]	AWE	100	1000	99.21(D)	Gives a PED-CED ear detection technique
Chowdhury et al. [18]	UND	302	942	98.2(I)	Detection using AdaBoost-based detector
Sarangi et al. [19]	UND E	114	464	96.6(I)	PHOG and LDA for feature extraction and dimensionality reduction
				3.395(EER)	
Alqaralleh and Toygar [20]	USTB-3	NA	NA	92.3(I)	Fusion of features from tragus and another part of the ear is done using LBP descriptor
Alshazly et al. [21]	IITD	221	793	97(I)	Gradient-based feature extraction methods are used



## 4 Conclusion

This paper features the use of computer vision and image processing technology in the field of Ear detection and recognition. It discusses various existing Ear identification, segmentation, detection, and recognition techniques, and approaches. It also provides a comparison of the performance of several approaches. Table 2 gives a comparative review of several existing techniques and their features. It shows the databases used by the author and the corresponding performances noted by them in terms of recognition rate, equal error rate, identification rate, and detection rate. Table 1 provides a comparative overview of the databases that are available for the researchers who wish to test or propose new Ear recognition and detection techniques. It includes features like number of subjects, number of images corresponding to the subjects; it also indicates whether the images are occluded or not.

## References

1. Burge, M., & Burger, W. (1997, May). Ear biometrics for machine vision. In *21st workshop of the Austrian association for pattern recognition* (pp. 275–282).
2. Moreno, B., Sanchez, A., & Vélez, J. F. (1999). On the use of outer ear images for personal identification in security applications. In *Proceedings IEEE 33rd Annual 1999 International Carnahan Conference on Security Technology* (pp. 469–476).
3. Mu, Z., Yuan, L., Xu, Z., Xi, D., & Qi, S. (2004). Shape and structural feature based ear recognition. *Advances in biometric person authentication* (pp. 663–670). Berlin: Springer.
4. Yuizono, T., Wang, Y., Satoh, K., & Nakayama, S. (2002, May). Study on individual recognition for ear images by using genetic local search. In *Proceedings of the 2002 Congress on Evolutionary Computation* (Vol. 1, pp. 237–242).
5. Burge, M., & Burger, W. (2000). Ear biometrics in computer vision. In *Proceedings 15th International Conference on Pattern Recognition* (Vol. 2, pp. 822–826).
6. Hurley, D. J., Nixon, M. S., & Carter, J. N. (2002). Force field energy functionals for image feature extraction. *Image and Vision Computing*, 20, 311–317.
7. Victor, B., Bowyer, K., & Sarkar, S. (2002). An evaluation of face and ear biometrics. In *Object recognition supported by user interaction for service robots* (Vol. 1, pp. 429–432).
8. Arbab-Zavar, B., & Nixon, M. S. (2011). On guided model-based analysis for ear biometrics. *Computer Vision and Image Understanding*, 115, 487–502.
9. Prakash, S., & Gupta, P. (2013). An efficient ear recognition technique invariant to illumination and pose. *Telecommunication Systems*, 52, 1435–1448.
10. Kumar, A., & Chan, T. S. T. (2013). Robust ear identification using sparse representation of local texture descriptors. *Pattern Recognition*, 46, 73–85.
11. Basit, A., & Shoaib, M. (2014). A human ear recognition method using nonlinear curvelet feature subspace. *International Journal of Computer Mathematics*, 91, 616–624.
12. Nigam, A., & Gupta, P. (2014, November). Robust ear recognition using gradient ordinal relationship pattern. In *Asian conference on computer vision* (pp. 617–632). Cham: Springer.
13. Pflug, A., Paul, P. N., & Busch, C. (2014, October). A comparative study on texture and surface descriptors for ear biometrics. In *2014 international carnahan conference on security technology* (pp. 1–6).
14. Anwar, A. S., Ghany, K. K. A., & Elmahdy, H. (2015). Human ear recognition using geometrical features extraction. *Procedia Computer Science*, 65, 529–537.

15. Youbi, Z., Boubchir, L., Bounneche, M. D., Ali-Chérif, A., & Boukrouche, A. (2016, June). Human ear recognition based on multi-scale local binary pattern descriptor and KL divergence. In *2016 39th international conference on telecommunications and signal processing* (pp. 685–688).
16. Ghoulmi, L., Draa, A., & Chikhi, S. (2016). An ear biometric system based on artificial bees and the scale invariant feature transform. *Expert Systems with Applications*, *57*, 49–61.
17. Emeršič, Ž., Gabriel, L. L., Štruc, V., & Peer, P. (2017). Pixel-wise ear detection with convolutional encoder-decoder networks. [arXiv:1702.00307](https://arxiv.org/abs/1702.00307).
18. Chowdhury, M., Islam, R., & Gao, J. (2017, June). Robust ear biometric recognition using neural network. In *2017 12th IEEE conference on industrial electronics and applications* (pp. 1855–1859).
19. Sarangi, P. P., Mishra, B. S. P., & Dehuri, S. (2017, February). Ear recognition using pyramid histogram of orientation gradients. In *2017 4th international conference on signal processing and integrated networks* (pp. 590–595).
20. Alqaralleh, E., & Toygar, Ö. (2018). Ear recognition based on fusion of ear and tragus under different challenges. *International Journal of Pattern Recognition and Artificial Intelligence*, *32*, 1856009.
21. Alshazly, H. A., Hassaballah, M., Ahmed, M., Ali, A. A. (2018, September). Ear biometric recognition using gradient-based feature descriptors. In *International conference on advanced intelligent systems and informatics* (pp. 435–445). Cham: Springer.
22. Emeršič, Ž., Štruc, V., & Peer, P. (2017). Ear recognition: More than a survey. *Neurocomputing*, *255*, 26–39.
23. West Virginia University Libraries. Retrieved January 15, 2019 from <https://lib.wvu.edu/databases/>.
24. Ear Recognition Laboratory Homepage at University of Science & Technology Beijing (USTB). Retrieved January 20, 2019 from <http://www1.ustb.edu.cn/resb/en/index.htm>.
25. UCR Library. Retrieved January 20, 2019 from <https://library.ucr.edu/research-services/databases>.
26. UND Chester Fritz Library. Retrieved January 15, 2018 from <https://library.und.edu/databases>.
27. UMIST Database. Retrieved January 20, 2019 from <https://www.sheffield.ac.uk/eee/research/iel/research/face>.
28. The XM2VTS Database. Retrieved January 17, 2019 from <http://www.ee.surrey.ac.uk/CVSSP/xm2vtsdb/>.
29. Face recognition technology (FERET). Retrieved January 15, 2019 from <https://www.nist.gov/programs-projects/face-recognition-technology-feret>.
30. The CAS-PEAL face database. Retrieved January 17, 2019 from <http://www.jdl.ac.cn/peal/top.htm>.
31. IIT Delhi Ear Database. Retrieved January 20, 2019 from [http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database\\_Ear.htm](http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Ear.htm).
32. IIT Kanpur Ear Database. Retrieved January 15, 2019 from <http://home.iitk.ac.in/~ganil/>.
33. AWE Database. Retrieved January 15, 2019 from <https://www.awedatabase.com/>.
34. UBEAR. Retrieved January 17, 2019 from <http://ubear.di.ubi.pt/>.
35. NCKU Database. Retrieved January 15, 2019 from [http://robotics.csie.ncku.edu.tw/Databases/FaceDetect\\_PoseEstimate.htm](http://robotics.csie.ncku.edu.tw/Databases/FaceDetect_PoseEstimate.htm).
36. YSU Database. Retrieved January 20, 2019 from <http://lib.ysu.am/libraries.html>.
37. Chang, K., Bowyer, K. W., Sarkar, S., & Victor, B. (2003). Comparison and combination of ear and face images in appearance-based biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*, 1160–1165.
38. Yuan, L., Mu, Z. C., Zhang, Y., & Liu, K. (2006, August). Ear recognition using improved non-negative matrix factorization. In *18th international conference on pattern recognition* (Vol. 4, pp. 501–504).
39. Choraś, M. (2008). Perspective methods of human identification: Ear biometrics. *Opto-Electronics Review*, *16*, 85–96.

40. Islam, S. M., Davies, R., Bennamoun, M., & Mian, A. S. (2011). Efficient detection and recognition of 3D ears. *International Journal of Computer Vision*, *95*, 52–73.
41. Lei, J., You, X., & Abdel-Mottaleb, M. (2016). Automatic ear landmark localization, segmentation, and pose classification in range images. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *46*, 165–176.

# Automatic Detection of Sleep Spindles Using Time Domain Features



Ghania Fatima, Omar Farooq and Shikha Singh

**Abstract** Sleep spindles are one of the unique rhythmic activities observed in sleep electroencephalogram (EEG). Detecting sleep spindles visually by sleep spindles is a difficult task as high skills and efforts are required. In this study, a methodology for detecting sleep spindles automatically has been proposed using band-pass filtering. Time domain features (energy and entropy) are used for classification. The extracted features have been used as inputs to Linear, Quadratic, and Mahalanobis classifier for spindle detection. Results show that the proposed method yields best results when using a Mahalanobis Classifier. The accuracy, sensitivity, and specificity recorded are 91.11%, 84.86%, and 89.73%, respectively. The sensitivity obtained in this study is more than most of the work done in sleep spindles detection using the same dataset.

**Keywords** Sleep spindles · EEG · Mahalanobis classifier

## 1 Introduction

Sleep is a primary function of the brain. A person's performance, physical movement, and learning capability are largely governed by his sleep pattern [1]. Sleep studies can improve our understanding about the mechanism of the brain. Since sleep spindles are one of the well-defined rhythmic activities present in the sleep EEG and a trademark of stage two of sleep, they are significant for brain research.

The sleep spindles are the transient EEG events which are unique to sleep. Their frequency was defined originally as 12–14 Hz in the Rechtschaffen and Kales (R&K) criteria [2]. It was increased to 11.75–16 Hz by Smith et al. [3], 11.5–15 Hz by Fish et al. [4], 10–16 Hz by Huupponen et al. [5], 11–16 Hz by Devuyst [6], and 11–15 Hz

---

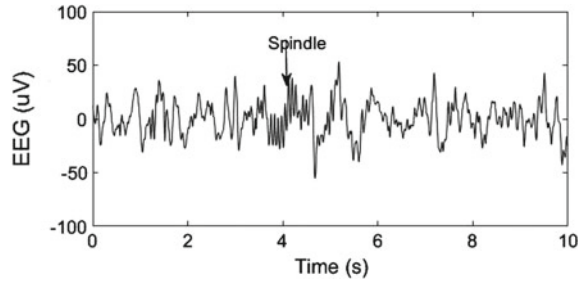
G. Fatima (✉) · O. Farooq · S. Singh  
Department of Electronics Engineering, Aligarh Muslim University, Aligarh, India  
e-mail: [ghaniafatima@gmail.com](mailto:ghaniafatima@gmail.com)

O. Farooq  
e-mail: [omar.farooq@amu.ac.in](mailto:omar.farooq@amu.ac.in)

S. Singh  
e-mail: [singhshikha65@gmail.com](mailto:singhshikha65@gmail.com)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_51](https://doi.org/10.1007/978-981-15-0694-9_51)

**Fig. 1** EEG recording (CZ-A1) with one spindle (10 s from excerpt 5)



by Durka [7]. According to the guidelines of American Academy of Sleep Medicine (AASM), the frequency range for sleep spindles is between 11–16 Hz [8]. Figure 1 shows an EEG recording with one spindle.

Studying about sleep spindles and its distribution and frequency of occurrence over a night's sleep is an important part of sleep research. Conventionally, overnight PSG recordings including Electroencephalogram (EEG) are scored visually by experts using R&K criteria [2]. However, visual scoring of sleep spindles is tedious and also has a high probability of human error. Therefore, several hardware and software methods for automatic detection of sleep spindles have been developed. Hardware methods comprise band-pass filter combined with systems performing frequency detection [3, 4]. Software methods consist of two main approaches. The first is band-pass filtering followed by level detection. Since there is inter-subject amplitude and frequency variability, recording specific amplitude threshold detection before performing level detection was proposed by Ray et al. [9] and Huupponen et al. [5]. The sensitivity of 98.96% and 73.50% and specificity of 88.49% and 98.50% respectively was reported. The second approach is feature extraction followed by decision-making for classification. Short time Fourier transform (STFT) and autoregressive (AR) modeling are common methods of feature extraction. Using STFT for feature extraction and linear discriminant analysis for classification, Anderer et al. [2] reported a sensitivity of 86% and specificity of 80%. Using STFT coefficients directly as classifier's inputs, Görür [10] reported an agreement rate of 88.70% with multilayer perceptron (MLP) and 95.40% with support vector machine (SVM). Using AR modeling, he achieved an average performance between 88.80 and 93.60% with MLP and between 93.30 and 96.00% with SVM.

In this study, a new method for automatic sleep spindles detection is presented in which time domain features are used for classification. The signal is first band-pass filtered before feature extraction and classification. Linear, Quadratic and Mahalanobis distance classifiers are used and their performances are compared.

## 2 EEG Dataset Used

The database used in this study of Sleep Spindles detection is the DREAMS Sleep Spindles Database of University of MONS-TCTS Laboratory (Stephanie Devuyst, Thierry Dutoit) and Université Libre de Bruxelles-CHU de Charleroi Sleep Laboratory (Myriam Kerkhofs). It is under terms of the Attribution-NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0) License. ([http://www.tcts.fpms.ac.be/~devuyst/Databases/DREAMS\\_databases\\_License.txt](http://www.tcts.fpms.ac.be/~devuyst/Databases/DREAMS_databases_License.txt)). Devuyst et al. proposed a standard assessment method for any sleep spindles detection algorithm and implemented it on their own detection procedure using this dataset [6]. The dataset was published on the internet for future research and performance comparison. The dataset, along with additional information is publicly available at <http://www.tcts.fpms.ac.be/~devuyst/Databases/DatabaseSpindles/>.

The data set consists of eight excerpts of central EEG channel (CZ-A1 and C3-A1). Each excerpt is a 30 min recording extracted from whole night PSG recordings of eight patients (four males and four females). Each excerpt is of 30 min and sampled at 200, 100, and 50 Hz. The recordings are stored in standard European Data Format (EDF). Two experts in sleep spindles have independently annotated the recordings. Since only the first six excerpts are annotated by both the experts, the last two excerpts are not used in this study.

## 3 Proposed Methodology

The three main steps in spindle detection procedure are preprocessing, feature extraction, and classification. The classification is done in two phases: training and testing. The system is first trained using features of known samples and then tested for its accuracy, sensitivity, and specificity. Once trained it is then used to detect spindles in unseen data.

### 3.1 Preprocessing

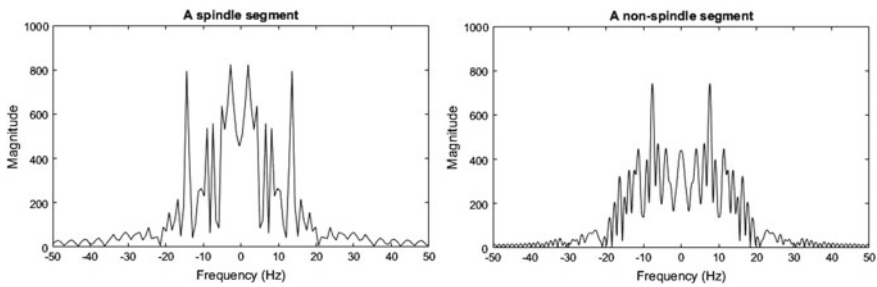
Preprocessing removes artifacts that appear in the EEG record. They are caused due to reasons such as the movement of head during signal recording, physical issues in electrode/lead/channel and connectivity issues between head and the device. Two types of noise are removed in the preprocessing. First is the power line noise which is 50 Hz line frequency interference and second is the baseline noise due to poor contact of electrodes and other physical problems. For this a notch filter and a band-pass filter is used respectively.

### 3.2 Feature Extraction and Classification

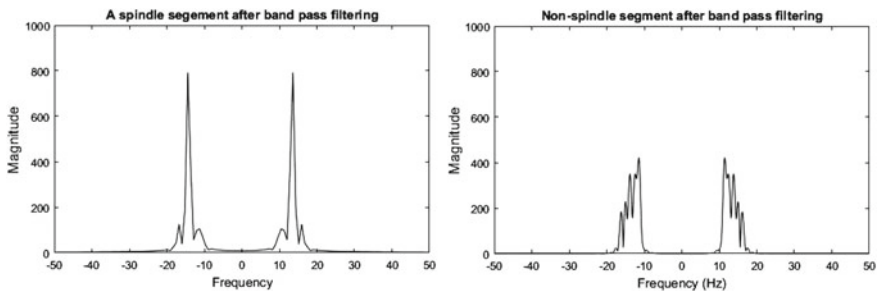
After preprocessing, the signal is segmented into 1 s epoch using a rectangular window technique. It is important to divide the EEG signal into smaller segments before analysis because EEG signals are nonstationary and many signal processing methods can be applied only on stationary signals. For small window durations, these signals behave like Quasi-stationary.

Each epoch is then band-pass filtered using an eighth order band-pass Butterworth filter with a frequency range of 11–16 Hz. Since the information of sleep spindles lies in this range (as shown in Fig. 2), it helps in retaining the frequency contents of only sleep spindles in each epoch and discarding all other frequencies (as shown in Fig. 3). The filtered segments are then used to calculate the features which give relevant information necessary for classification.

The features used in the proposed methodology are energy of the signal and Shannon entropy. These features when calculated after band-pass filtering make the two classes (spindles and non-spindles) separable.



**Fig. 2** Frequency spectrum of a sleep spindle segment and non-spindle segment demonstrating that sleep spindles have a large magnitude in the frequency range of 11–16 Hz



**Fig. 3** Frequency spectrums of the segments after band-pass filtering them in the spindle frequency range (11–16 Hz)

- The energy of a signal gives a measure of the strength of a signal. An EEG signal values are both positive and negative and, therefore, there are two options: computation of area under the square of the function or calculating area under the absolute value of the function. The first choice is preferred because of its mathematical tractability and similarity to Euclidean.
- Entropy is a way to quantify the amount of uncertainty or randomness in the pattern which is also roughly equivalent to the amount of information contained in the signal. In the proposed methodology, total wavelet entropy is calculated where wavelets are used to decompose the EEG signal into multiple resolution levels.

The last step in the detection algorithm is to classify the extracted features (energy and entropy) into the two different classes for detecting sleep spindles. The extracted features (energy and entropy) from each segment are used as input to Linear, Quadratic and Mahalanobis distance classifier and a comparison of the performances of these classifiers are made.

### 3.3 Performance Parameters

The performance of the proposed method is evaluated with the help of three parameters: accuracy, sensitivity, and specificity.

- **Accuracy:** Accuracy gives the number of correct classifications. It is the ratio of true results (positive and negative) to the total number of cases. It is defined as:

$$A_{cc} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Where, TP (true positive) is the number of sleep spindle segments correctly detected, FN (false negative) is the number of sleep spindles that are wrongly identified as non sleep spindles, TN (true negative) is the number of non sleep spindles that are correctly identified and FP (false positive) is the number of non sleep spindles that are incorrectly identified as spindles.

- **Sensitivity:** It is the ratio of true positive prediction to the total number of positive cases and is defined as:

$$S_e = \frac{TP}{TP + FN} \quad (2)$$



- **Specificity:** Specificity is the ratio of true negative prediction to the total number of negative cases and is defined as:

$$S_c = \frac{TN}{TN + FP} \quad (3)$$

## 4 Results and Discussions

Experiments were done on the sleep dataset defined in Sect. 3 of the paper using the proposed method and the performance was evaluated. Three different experiments were performed. In the first experiment, training and testing of the system were done for each subject individually and then the average values of the performance parameters were taken. In the second experiment, the system was trained using 75% of the data for all subjects taken together and testing was done for the remaining 25% of the data. These two experiments are subject dependent classification of sleep spindles. In the third experiment, the system was trained using the data of four subjects and tested for remaining two subjects. This was done using all possible 15 combinations of subjects and the average values were taken. This is subject independent classification of sleep spindles. The classification is done using three different classifiers: Linear, Quadratic, and Mahalanobis and their results are compared.

### 4.1 Subject Dependent Classification

Table 1 gives the results of the proposed method when classification for each subject is done individually and the average values of the performance parameters are calculated.

Table 2 gives the results of the proposed method when the system is trained using 75% of data of all subjects taken together and testing for the remaining 25% data. A four-fold cross-validation is performed.

From Tables 1 and 2, we observe that the Linear classifier gives high accuracy and specificity but its sensitivity is low. Whereas, the Quadratic classifier gives high

**Table 1** Performance comparison of the classifiers when the sleep spindles detection for each subject is done individually

Performance parameters	Linear (%)	Quadratic (%)	Mahalanobis (%)
Accuracy	90.97	78.70	89.46
Specificity	82.83	91.81	91.41
Sensitivity	91.45	78.25	89.37

**Table 2** Performance comparison when detection of sleep spindles is done by training the system with 75% of the data of all subjects taken together and testing for the remaining 25% data

Performance parameters	Linear (%)	Quadratic (%)	Mahalanobis (%)
Accuracy	90.07	67.30	83.81
Specificity	82.37	93.04	88.44
Sensitivity	90.47	66.01	83.58

**Table 3** Performance of different classifiers when subject independent classification of spindles and non-spindles is done

Performance parameters	Linear (%)	Quadratic (%)	Mahalanobis (%)
Accuracy	94.14	84.65	91.11
Specificity	77.59	88.88	84.86
Sensitivity	94.89	84.35	89.73

sensitivity but the accuracy and specificity are not very good. Mahalanobis classifier gives a decent performance in terms of accuracy, sensitivity, and specificity.

## 4.2 Subject Independent Classification

Table 3 gives the performance comparison of the three classifiers when the classification is subject independent. Here the data of four subjects are used as inputs to the classifier for training and its performance is tested for the remaining two unseen subjects. All possible 15 combinations of the data of six subjects are tested and the average values of the results are taken.

Comparison of the result obtained in this study is made with some of the previous studies on sleep spindles detection for better evaluation of the performance (as shown in Table 4). The studies selected for comparison have been done on the same dataset as described in Sect. 3. Devuyst et al. presented a systematic assessment method of sleep spindles detection. A window size 0.5 s with a 20% overlap were used. The sensitivity and specificity reported were 70.20 and 98.60% respectively [6]. Imtiyaz et al. used Teager Energy and Spectral edge frequency as features for sleep spindle detection. A window size of 0.25 s and a 50% overlapping were used. The accuracy, sensitivity and specificity reported were 91%, 80.00% and 98.00%, respectively [11]. Nonclercq et al. used amplitude-based features for detection. A window of 0.5 s duration with 50% overlap was used. They reported a sensitivity of 78.50% and specificity of 94.20% [12]. Patti et al. identified sleep spindles using Gaussian mixture model. A 1.5 s window was used. The sensitivity of 74.90% was obtained [13]. Parekh et al. [14] did the detection based on oscillatory low frequency components. They obtained an accuracy of 96.4%, sensitivity of 70.00%, and specificity of 97.80%. Tsanas et al. did the detection using continuous wavelet transform and local weighted smoothing for sleep spindles detection and obtained a sensitivity of 76.00% and specificity 92.00%

**Table 4** Comparison of the result obtained using the proposed methodology with previous works done on sleep spindle detection using the same dataset. (L, Q, and M stands for Linear, Quadratic and Mahalanobis classifiers respectively)

Author	Accuracy (%)	Sensitivity (%)	Specificity (%)
Devuyt et al. [6]	–	70.20	98.6
Imtiyaz et al. [11]	91.00	80.00	98.00
Nonclercq et al. [12]	–	78.50	94.20
Patti et al. [13]	–	74.9	–
Parekh et al. [14]	96.40	70.00	97.80
Tsanas and Clifford [15]	–	76.00	92.00
Zhuang et al. [16]	–	50.98	99.00
Zhou et al. [17]	–	70.70	96.30
Proposed method	94.14 (L) 84.65 (Q) 91.11 (M)	77.59 (L) 88.88 (Q) 84.86 (M)	94.89 (L) 84.35 (Q) 89.73 (M)

[15]. Zhuang and Peng detected sleep spindles using a sliding window probability estimation method. The average sensitivity of 50.98% and specificity of 99.00% were recorded [16]. The proposed method achieves better sensitivity than most of the previous studies done on this dataset. The accuracy and specificity are less but comparable.

## 5 Conclusions

Sleep spindles are useful for classification of sleep stages and are related to brain maturation. Their distribution in the sleep EEG can be used to describe the morphology of the sleep EEG. Detecting sleep spindles visually in a full night’s recording is time-consuming and also tiring. An automated sleep spindle detection system eliminates the subjectivity in the detection and also reduces the workload of the expert.

In this study, a new methodology of automatic sleep spindle detection is proposed. Comparative study of results for three different classifiers (Linear, Quadratic and Mahalanobis) is done. It is found that for the technique and features used, a Linear classifier gives high accuracy and specificity but the sensitivity is low. But if we use a Mahalanobis classifier, there’s significant improvement in sensitivity without much degradation in accuracy and specificity. The accuracy, sensitivity, and specificity as reported in this study using a Mahalanobis classifier is 91.11%, 84.86%, and 89.73%,

respectively. The sensitivity achieved is higher than that achieved in most of the previous works. In the end, a comparison of the results obtained using the proposed algorithm is made with that of the previous works done using the same database (DREAMS sleep database). In previous works, a tradeoff between sensitivity and specificity is found (higher specificities are achieved but the sensitivity obtained is low). Here, we have tried to increase the sensitivity achieved without compromising with the accuracy and specificity much.

**Acknowledgements** The authors would like to extend their gratitude to the University of MONS-TCTS Laboratory (Stephanie Devuyst, Thierry Dutoit) and Université Libre de Bruxelles-CHU de Charleroi Sleep Laboratory (Myriam Kerkhofs) for making the DREAMS dataset available publicly for research.

## References

1. Huang, C. S., Lin, C. L., Ko, L. W., Liu, S. Y., Sua, T. P., & Lin, C. T. (2013). A hierarchical classification system for sleep stage scoring via forehead EEG signals. In *Proceeding of 2013 IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind and Brain, CCMB 2013—2013 IEEE Symposium Series on Computer Intelligence SSCI* (pp. 1–5).
2. Anderer, P., Gruber, G., Parapatics, S., & Dorffner, G. (2007). Automatic sleep classification according to Rechtschaffen and Kales. In *Proceedings of Annual International Conference IEEE Engineering in Medicine and Biology* (pp. 3994–3997).
3. Smith, J. R., Funke, W. F., Yeo, W. C., & Ambuehl, R. A. (1975). Detection of human sleep EEG waveforms. *Electroencephalography and Clinical Neurophysiology*, *38*, 435–437.
4. Fish, D. R., Allen, P. J., & Blackie, J. D. (1988). A new method for the quantitative analysis of sleep spindles during continuous overnight EEG recordings. *Journal of Sleep Research*, *70*, 273–277.
5. Huupponen, E., Värri, A., Himanen, S. L., Hasan, J., Lehtokangas, M., & Saarinen, J. (2000). Optimization of sigma amplitude threshold in sleep spindle detection. *Journal of Sleep Research*, *9*(4), 327–334.
6. Devuyst, S., Dutoit, T., Stenuit, P., & Kerkhofs, M. (2011). Automatic sleep spindles detection—overview and development of a standard proposal assessment method. In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 1713–1716). IEEE.
7. Durka, P. J., & Blinowska, K. J. (1995). Analysis of EEG transients by means of matching pursuit. *Annals of Biomedical Engineering*, *23*(5), 608–611.
8. Iber, C., Ancoli-Israel, S., Chesson, A., & Quan, S. F. (2007). *The AASM manual for the scoring of sleep and associated events: Rules, terminology and technical specifications*. Westchester, Illinois (IL): American Academy of Sleep Medicine.
9. Ray, L. B., Fogel, S. M., Smith, C. T., & Peters, K. R. (2010). Validating an automated sleep spindle detection algorithm using an individualized approach. *Journal of Sleep Research*, *19*(2), 374–378.
10. Görür, D. (2003). Automated detection of sleep spindles, MSc thesis, Middle East Technical University.
11. Imtiyaz, S. A., Saremi-Yarahmadi, S., & Rodriguez-Villegas, E.: Automatic detection of sleep spindles using Teager energy and spectral edge frequency, *IEEE Biomedical Circuits and Systems Conference BioCAS 2013*, 262–265.
12. Nonclerq, A., Urbain, C., Verheulpen, D., Decaestecker, C., Bogaert, P. V., & Peigneux, P. (2013). Sleep spindle detection through amplitude–frequency normal modelling. *Journal of Neuroscience Methods*, *214*, 192–203.

13. Patti, C. R., Chaparro-Vargas, R., & Cvetkovic, D. (2014). Automated sleep spindle detection using novel EEG features and mixture models. In *2014 36th Annual International Conference IEEE Engineering Medicine Biology Society EMBC* (Vol. 1, pp. 2221–2224).
14. Parekh, A., Selesnick, I. W., Rapoport, D. M., & Ayappa, I. (2015). Detection of K-complexes and sleep spindles (DETOKS) using sparse optimization. *Journal of Neuroscience Methods*, *251*, 37–46.
15. Tsanas, A., & Clifford, G. D. (2015). Stage-independent, single lead EEG sleep spindle detection using the continuous wavelet transform and local weighted smoothing. *Frontiers Human Neuroscience*, *9*, 181.
16. Zhuang, X., Li, Y., & Peng, N. (2016). Enhanced automatic sleep spindle detection: A sliding window based wavelet analysis and comparison using a proposal assessment method. *Applied Informatics*.
17. Zhou, S., Zhang, X., & Yu, Z. (2017). A sleep spindle detection algorithm based on SVM and WT. In: *2017 29th Chinese Control and Decision Conference (CCDC)*, (pp. 2213–2217). Chongqing.

# A Review on Lung and Nodule Segmentation Techniques



Bhawana Kamble, Satya Prakash Sahu and Rajesh Doriya

**Abstract** Computer Aided Diagnosis (CAD) systems for automatic detection of pulmonary diseases and lung cancer mainly depend on the segmentation of different pulmonary components like right and left lung lobes, airways, vessels, and nodules from the medical imaging modalities like CTs, MRIs, etc. Lung segmentation and nodule segmentation are the important steps to detect any lung related abnormalities. It requires many image processing operations to be performed on the medical images. Computed Tomography (CT) imaging is the most preferred modal because of its popularity, ease of use, and capability of showing different anatomical structures of thorax region. This review paper includes a study of various state of the art techniques explaining the methods applied on CT scans to find the ROIs along with their segmentation accuracies parameters in terms of similarity coefficient, mean error, and overlap ratio.

**Keywords** CAD system · Lung segmentation · Region growing · Thresholding · Nodule segmentation

## 1 Introduction

According to the latest WHO research in 2017, Lung Disease Deaths in India reached 896,779 or 10.19% of total deaths. The age adjusted death rate is 96.92 per 100,000 of population ranks India 4 in the world [1]. Lung diseases are one of the leading death rates in the world. According to the Surveillance, Epidemiology, and End Results (SEER) program which dispense the information about the cancer facts and figures

---

B. Kamble (✉) · S. P. Sahu · R. Doriya  
Department of Information Technology, National Institute of Technology, Raipur (C.G), India  
e-mail: [bkamble.mtech2017.it@nitrr.ac.in](mailto:bkamble.mtech2017.it@nitrr.ac.in)

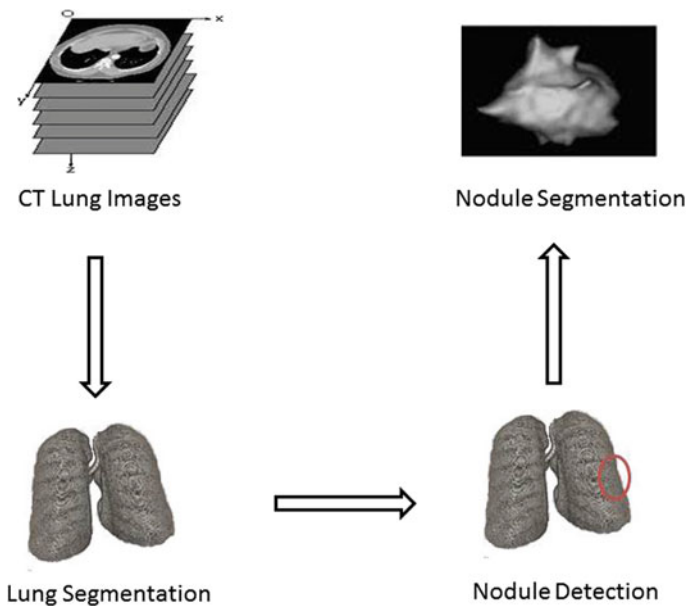
S. P. Sahu  
e-mail: [spsahu.it@nitrr.ac.in](mailto:spsahu.it@nitrr.ac.in)

R. Doriya  
e-mail: [rajdor.it@nitrr.ac.in](mailto:rajdor.it@nitrr.ac.in)

in the U.S, the survival rate has increased to 18.6% from 2008 to 2014 [2] and it can be further increased if this type of diseases is treated in the early stages. Nowadays the detection of the diseases is done by radiologists by using CT scan images and it is very difficult to identify the pattern of the disease. The main computational systems developed to benefit radiologists is CADe (computer aided detection system) which have some goals such as, detection is done in early stages, better accuracy in diagnosis, and should take less time to evaluate the results. For accurate and correct result some techniques are introduced by the following papers. These papers describe the working of segmentation techniques and obtain accurate results when compared to ground truths.

## 2 Lung Segmentation Techniques

Lung Segmentation is a computer based process to obtain the boundary of the lungs from thoracic on Computed Tomography (CT) images. Several algorithms such as region growing, threshold techniques, etc. are applied to the medical images to perform segmentation. The Fig. 1 describes the process of lung and nodule segmentation in a typical CAD system. The input of this CAD system is the medical images which are obtained using a suitable mode. The first step is a preprocessing step which may



**Fig. 1** Flow of lung and nodule segmentation process in a typical Computer Aided Diagnosis system

contain transformation and image format specific operations. The lung segmentation is an essential step, it basically extracts the portion of lungs from CT chest image for finding the accurate areas of interest. Lung segmentation is done to lessen the search space and decrease the overhead for the next process which is lung nodules detection. Nodule detection is done to spot the locations of lung nodules then the nodules are segmented which has a set of features, like shape, volume, etc., that are extracted and used for the diagnosis [3]. The following are some techniques used for lung and nodule segmentation.

## ***2.1 3D Region Growing***

Da Nobrega et al. [4] used Insight Toolkit (ITK) segmentation and Visualization Toolkit (VTK) for the representation of the segmented region in three dimensions using volume rendering techniques. CT images of the pulmonary region of the chest are taken as input and segmentation is done as the first step on the basis of correction between atomic number and radiographic density of the given material. A threshold point is applied for segmenting the normal aerated region ( $-900$  to  $-500$  HU) and hyper inflated region ( $-1000$  to  $-900$  HU). Then the region growing algorithm is applied to the internal connection point to extract the lung segmentation. This method came with the new approach of the lung segmentation with the assistance of several ITK techniques and it produces result of 98.76% which is more accurate in comparison with AOSP (97.45%). Yim et al. [5] proposed a method to identify lung segmented area in lung CT images. Lungs and aviation routes are extricated by connected component labeling and inverse seeded region growing. Then, large airways and trachea are outlined from the lungs by 3D region growing. At last, exact lung region borders are gotten by subtracting the output of the second step from that of the initial step.

## ***2.2 Nonnegative Matrix Factorization (NMF)***

Hosseini-asl et al. [6] proposed a framework to model both the spatial interaction and first-order visual appearance of the lung based on a new NMF method. For an efficient segmentation of the 3D lung, again Hosseini-asl et al. [7] proposed an incremental constrained nonnegative matrix factorization (ICNMF). Here the idea of Constrained NMF and incremental NMF are combined to form ICNMF. 3D lung segmentation is done on CT images by 3D region growing method. This proposed method is applied on both synthetic and vivo data using three performance matrices: Dice Similarity Coefficient (DSC), Absolute Lung Volume Difference (ALVD), and Modified 95percentile Hausdorff Distance (MHD). The ICNMF discloses the robust features for encoding the voxel neighbor with the smooth descriptor.



### **2.3 Adaptive Crisp Active Contours**

The proposed method of Rebouças Filho et al. [8] has reduced the analysis time and increased the accuracy. Machine learning techniques with active contour combined with the work of Sun et al. [9], Mansoor et al. [10], Wei et al. [11] are applied to segment the lung region in 3D effectively. The method moves points of the model using information from model shape and image voxel. By minimizing the energy of the 3D adaptive crisp active contour method, one point is moved which is calculated by the combination of 3D adaptive internal energy and 3D adaptive crisp external energy. The steps in segmentation starts with opening DICOM images then external energy is calculated. Then a 3D initializing voxel is extracted and then the iterative method of 3D adaptive crisp ACM is applied in order to decrease the energy by moving points. The iteration stops when the volume does not increase. GLU library is used to view the segment in 3D. Sarmiento et al. [12] emphasis on segmentation and reconstruction of CT images of lungs by using adaptive crisp active contour model. The analysis is done in 3D by using OpenGL library.

### **2.4 Automatic Lung Segmentation**

Silva et al. [13] presents an automated method for identifying lungs in CT images. To get lung region, a threshold with cutoff value ranging  $-900-0$ HU is taken which includes the possible intensity of lungs. For removing airways the algorithm looks for the trachea and by using a new threshold value trachea is detected and isolated using a morphological operation like dilation. Next, the right and the left lung is separated, for which sequential erosion is performed. The inferior and superior border of lungs are obtained then the morphological operation is applied to hold the maximum area of the lung and at last, subtraction is done between the original image and the border obtained. A method is proposed by Rikxoort et al. [14] which is a hybrid method to segment lung region. This method consists of described steps: 1. by using an automatic 3D algorithm, region growing and morphological operation lung field is segmented. 2. Automatic error detection. 3. At last multiatlas segmentation is performed. Noor et al. [15] proposed another automatic lung segmentation strategy dependent on the local and global system. The local framework depends on the surface and the global framework depends on the morphology with an installed control feedback that identifies and corrects substantial deviations and nonsuccess of segmentation.

### **2.5 MGRF**

Soliman et al. [16] proposed validation approach for validating the segmentation method which generates 3D phantom to validate segment. Joint Markov–Gibbs

Model is used to describe the original 3D realist phantom and filtered GSS. Bayesian fusion is used to fuse segmented result. Abdollahi et al. [17] used joint Markov-Gibbs random field to describe the map of regions. The border regions are specified through the estimation of signal distribution by Linear Combination of Discrete Gaussians (LCDG). The beginning segmentation from the native and the generated Gaussian Scale Space (GSS) CT images are based on the LCDG models, next, they are iteratively cultured using an MGRF model. Then initial segmentations are merged together using a Bayesian fusion approach to obtain the final segmentation of the lung area. Table 1 shows the research studies for lung segmentation by various authors with datasets, adopted methods and achieved performance.

### 3 Nodule Segmentation and Classification

Lung nodules are round in shape and look like a white shadow in CT images. Nodules can be categorized into four different categories: 1. Juxtapleural, which is attached to the wall of parenchyma. 2. Juxta-vascular, nodules are attached to blood vessels. 3. Ground-Glass Opaque (GGO), nodules which are sub solid in nature [22]. Lung nodules can be noncancerous (benign) or cancerous (malignant). These nodules have different shapes and HU values so it is not efficient to segment nodules with the same method. Chen et al. [23] proposed a method for juxta-vascular nodules. Kuhnigk et al. [24] presented a method to perform rough segmentation of juxtapleural nodule. Several methods or techniques are included in this review like Gaussian Filtering, SVM, 3D active contour, Hessian based approaches, and so forth. Below are some techniques to segment lung nodule.

#### 3.1 *Gaussian Filtering Regularized Level Set*

Sudipta [25] proposed a framework for all nodules dependent on the inside surface (nonsolid, part solid/solid) and outer connection (juxta-vascular and juxta-pleural). In the proposed system, first nodules are ordered into nonsolid, part solid/solid then two separate division techniques are produced for nonsolid, part solid/solid nodules, individually. Wang et al. [26] designed a framework to segment two types of lung nodules solid and nonsolid in 3D. The method based on Selective binary and Gaussian Filtering Regularized Level Set (SBGFRLS) is applicable for segmentation of both types of nodules. The framework has following steps: 1. Preprocessing, where the lesions attached with wall are removed by thresholding followed by hole filling algorithm and blood vessels are removed by edge scrap removal method. 2. Rough Segmentation is done by SBGFRLS to create initial contour. 3. GAC refined segmentation, Geodesic Active Contour [27] is an edge based active contour model and has the ability to locate both the part solid and solid lesions. A 3D visualization using VTK is formed and volume is counted.

**Table 1** Comparison of different techniques used in lung segmentation

Author, Year	Dataset	Image Size	Technique	Result
Sahu et al. [18]	20 CT scan images including 10 CT juxtapleural cases	512 × 512	FCM, thresholding, morphological operation	Overlap ratio = 99.94% DCS = 0.971 JI = 0.944
Da Nobrega et al. [4]	30 CT scan images build by [8]	NA	Thresholding, 3D region growing	S = 98.09% Sp = 99.87% (By 3D RG)
Rebouças Filho et al. [8]	CT images	512 × 512	3D adaptive crisp active contour	Fm = 99.22%
Soliman et al. [19]	3D CT images from LOLA11, ISBI	512 × 512 × 270 – 450	Adaptive shape modeling	Accuracy = 98.0 ± 1.0%
Ng et al. [20]	HRCT images	512 × 512	Otsu threshold	DCS = 98.32%
Hosseini-Asl et al. [7]	17 CT images LOLA 11	512 × 512 × 390 voxel	Incremented constrained non negative matrix factorization	DSC = 0.96 ± 0.01 ALVD = 0.87 ± 0.62 MHD = 9.0 ± 0.001
Hosseini-Asl et al. [6]	CT images	512 × 512 × 390	Non negative matrix factorization	DSC = 0.966
Soliman et al. [16]	CT images	512 × 512 × 390	3D GGMRF	DSC = 0.9939
Silva et al. [13]	CT images using Philips Gemini GXL PET-CT scanner	512 × 512 pixel	Gaussian filter, thresholding, region growing, morphological operation	Ds = 2.47% and 0.078%
Abdollahi et al. [17]	CT images	512 × 512 × 390	GSS generation, 3D Joint MGRF	DSC (mean) = 0.960
Ren et al. [21]	CT images	512 × 512	Adaptive threshold, 3D region growing	Accuracy = 91.55
Rikxoort et al. [14]	50 scans from LIDC 50 scans from ILD 50 scans NELSON	0.75 mm pixel size 0.49–0.75 mm 0.531–0.836 mm	Region growing, morphological smoothing, multiatlas segmentation, automatic error detection	Volumetric overlap = 0.95 Hausdorff distance = 23.65 mm
Yim et al. [5]	CT images	512 × 512 pixel	3D Region growing, morphological operation	Difference of 1.2 pixels compared to ground truth

Abbreviations: Sp: Specificity =  $TN/(TN + FP)$ ; S: Sensitivity =  $TP/(FN + TP)$ ; 3D RG = 3 dimension Region growing; AOSP = Automatic Segmentation with Osirix Software plugin; DSC = Dice similarity coefficient; ALVD = Absolute lung volume difference; MHD = Modified 95-percentile Hausdorff distance; LOLA11 = Loband Lung Analysis 2011; Fm; F-measure =  $2*((recall*precision)/(recall + precision))$ ; Ds =  $((Aalgo - Agold \text{ standard})/Aalgo)\%$ ; ISBI = International Symposium on Biomedical Imaging; JI = Jaccard index; HRCT = High Resolution CT

### ***3.2 Gaussian-Mixture Model, Tsallis Entropy, SVM, Active Contour***

Santos et al. [28] proposed a method for automatic detection of the lung nodules (2–10 mm) by pattern recognition and image processing techniques. The process starts by taking CT images as input then nodule is segmented by threshold segmentation, region growing, rolling ball technique, and Gaussian-mixture model. Hessian Matrix is used for internal structure segmentation. Then the spherical structure is detected by applying Support Vector Machine (SVM), Tsallis entropy, and Shannon entropy. Then a 3D visualization of the detected nodule is shown. Nithila et al. [29] presented a region based active contour model for the reconstruction of lung area using selective binary and gaussian filtering with new signed pressure force function (SBGF-new SPF) and clustering techniques named Fuzzy C-Means (FCM) for nodule segmentation. Hao et al. [30] presented an automatic segmentation of juxta-vascular nodules which is based on Local binary fitting(LBF) active contour model with the joint vector and Standard uptake value (SUV) information entropy for evolution of contour curve to end at the edge of the nodule correctly.

### ***3.3 Region Based, Region Growing***

Chen et al. [31] developed a novel region based method for segmenting lung and lung nodules using chest CT images. The purpose of this method is to detect the boundary, volume, and position of the nodule. Segmentation process starts with the preprocessing step which includes noise filtering then the lung segmentation is done by using an absolute differentiation method. Region growing is applied to obtain the tumors and lung area. A 3D structure grid method is applied to reconstruct the three-dimensional volume for more reliable vision. Dehmeshki et al. [32] presented a region growing method for segmentation of lung nodules. A fusion of peripheral contrast as the halting criterion, intensity information and distance as the growing mechanism and fuzzy connectivity is used.

### ***3.4 3D Shape Analysis, SVM***

Oseas et al. [33] proposed a methodology to classify nodule and non-nodule by using pattern recognition and image processing techniques. Feature extraction is done on the basis of features like shape diagram and proportion of measures. Cylindrical based analysis is done to analyze the shape. To test the proposed method SVM is used. It has some limitation due to the similar property of some structures which fall in the same classes. Zhu et al. [34] proposed a method to evaluate the result obtained from SVM based classifier to differentiate between benign and malignant pulmonary

**Table 2** Comparison of different techniques used in nodule segmentation

Author, Year	Dataset	Features used classification	Technique	Result
Muhammad et al. [36]	888 CT images LIDC	A hybrid geometric texture feature	FODPSO based optimal Thresholding, geometric fit in parametric form	S = 95.6%, Sp = 97%
Oseas et al. [33]	833 CT images LIDC	Cylinder based analysis	SVM	Accuracy = 95.33%
John et al. [37]	10 CT scan from LIDC and PLD	Morphological and intensity features: Area, solidity, eccentricity	Multilevel thresholding	No. of nodules calculated
Rendon-Gonzalez et al. [38]	CT images	Area, eccentricity, circularity	SVM	S = 84.93% Sp = 80.9%
Zhou et al. [39]	CT images	Shape, intensity, texture	SVM	S = 99% Sp = 98.66%
Wang et al. [26]	280 group of solid, 80 nonsolid CT scan images (TCIA)	Geodesic active contour	Selective binary and gaussian filtering regularized Level set method	Avg. deviation = 5.46%(solid) 11.06%(nonsolid)
Santos et al. [28]	140 CT images LIDC	Tsallis entropy Shannon entropy	Gaussian mixture, SVM	S = 90.6%, Sp = 85%
Chen et al. [31]	520 DICOM images from China medical hospital, Taiwan	NA	Median filtering, Absolute Differentiation, 3D structure grid	3D view of nodule
Schilham et al. [40]	Images from JSRT	Position, detector feature	Gaussian scale space techniques	Detection rate = 67% (close to ground truth)

Abbreviations: Sp: Specificity =  $TN/(TN+FP)$ ; S: Sensitivity =  $TP/(FN+TP)$ ; LIDC = Lung Image Database Consortium; DSC = Dice similarity coefficient; TCIA = The Cancer Imaging Archive; SVM = Support Vector Machine; Accuracy =  $((TP+TN)/ TP+TN+FN+FP)$ ; JSRT = Japanese Society of Radiological Technology

nodule. Feature selection is done by using the genetic algorithm to subset group on the basis of different features like shape, size, etc. Diciotti et al. [35] presents a technique which depends on a neighborhood shape analysis of the initial division making utilization of 3-D geodesic distance map. The improvement strategy has the profit position that it locally refines the nodule division along with perceived vessel connections only, without changing the nodule edge. Table 2 shows the research studies for nodule segmentation by various authors with datasets, feature analysis, classification methods and achieved performance.

## 4 Discussion and Conclusion

The principal motive of this paper is to study the various proposed lung and nodule segmentation techniques. This will help to know how operations are performed on CT images and what we can enhance in the next model to achieve more accuracy. The studied papers used techniques like Thresholding, SVM, Region growing, and 3D shape analysis. The literature study reveals that the results of existing methods may be enhanced with hybrid techniques. Further the level of automation is required to be increased in future so as to be used in clinical practices to assist the radiologists or medical experts.

## References

1. Lung Disease in India. 2017. <https://www.worldlifeexpectancy.com/india-lung-disease>.
2. NCI online. (2016). Lung and Bronchus Cancer—Cancer Stat Facts, SEER Stat Fact Sheets. Lung and Bronchus Cancer.
3. El-Baz, A., et al. (2013). Computer-aided diagnosis systems for lung cancer: Challenges and methodologies. *International Journal of Biomedical Imaging*.
4. Da Nobrega, R. V. M., Rodrigues, M. B., & Filho, P. P. R. (2017, June). Segmentation and visualization of the lungs in three dimensions using 3D region growing and visualization toolkit in CT examinations of the chest. In *Proceedings of IEEE Symposium on Computer based Medical System* (Vol. 2017, pp. 397–402).
5. Yim, Y., Hong, H., & Shin, Y. G. (2005). Hybrid lung segmentation in chest CT images for computer-aided diagnosis. In *Proceedings of 7th International Workshop on Enterprise Networking and Computing in Healthcare Industry Healthcom 2005* (pp. 378–383).
6. Hosseini-asl, E., Zurada, J. M., & El-baz, A. (2014). Lung segmentation based on non-negative matrix factorization. Electrical and Computer Engineering Department, University of Louisville, Louisville, KY, USA. Bioengineering Department, University of Louisville, Louisville, KY, USA. Information Tech, no. 502 (pp. 877–881).
7. Hosseini-Asl, E., Zurada, J. M., Gimel-farb, G., & El-Baz, A. (2016). 3-D lung segmentation by incremental constrained nonnegative matrix factorization. *IEEE Transactions on Biomedical Engineering*, 63(5), 952–963.
8. Reboucas Filho, P. P., Cortez, P. C., da Silva Barros, A. C., Victor, V. H., & Tavares, R. S. J. M. (2017). Novel and powerful 3D adaptive crisp active contour method applied in the segmentation of CT lung images. *Medical Image Analysis*, 35, 503–516.
9. Sun, S., Bauer, C., & Beichel, R. (2012). Automated 3-D segmentation of lungs with lung cancer in CT data using a novel robust active shape model approach. *IEEE Transactions on Medical Imaging*, 31(2), 449–460.
10. Mansoor, A., et al. (2014). *Lung Segmentation*, 33(12), 2293–2310.
11. Wei, J., & Li, G. (2014). Automated lung segmentation and image quality assessment for clinical 3-D/4-D-computed tomography. *IEEE Journal of Translational Engineering in Health and Medicine*, 2.
12. PedrosaReboucasFilho, P., Sarmiento, R. M., Cortez, P. C., Carlos da Silva Barros, A., Hugo, V., & de Albuquerque, C. (2015). Adaptive crisp active contour method for segmentation and reconstruction of 3D lung structures. *International Journal of Computer Applications*, 111(4), 1–8.
13. Silva, S., Ferreira, N. C., & Caramelo, F. (2012). Dataset: 3D Automatic lung segmentation in low-dose CT (pp. 2–5).

14. Van Rikxoort, E. M., De Hoop, B., Viergever, M. A., Prokop, M., & Van Ginneken, B. (2009). Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection. *Medical Physics*, 36(7), 2934–2947.
15. Noor, N. M., Than, J. C. M., Rijal, O. M., Anzidei, M., Saba, L., & Suri, J. S. (2015). Automatic lung segmentation using control feedback system: Morphology and texture paradigm.
16. Soliman, A., Khalifa, F., Alansary, A., Gimel'Farb, G., & El-Baz, A. (2013). Segmentation of lung region based on using parallel implementation of joint MGRF: Validation on 3D realistic lung phantoms. In *Proceedings of International Symposium on Biomedical Imaging* (pp. 864–867).
17. Abdollahi, B., Soliman, A., Civelek, A. C., Li, X. F., Gimel'Farb, G., & El-Baz, A. (2012). A novel gaussian scale space-based joint MGRF framework for precise lung segmentation. In *Proceeding of International Conference on Image Processing ICIP* (pp. 2029–2032).
18. Sahu, S. P., Agrawal, P., Londhe, N. D., & Verma, S. (2017). A new hybrid approach using fuzzy clustering and morphological operations for lung segmentation in thoracic CT images. *Biomedical and Pharmacology Journal*, 10(4), 1949–1961.
19. Soliman, A., et al. (2017). Accurate lungs segmentation on CT chest images by adaptive appearance-guided shape modeling. *IEEE Transactions on Medical Imaging*, 36(1), 263–276.
20. Ng, C. R., et al. (2017). Preliminary 3D performance evaluation on automatic lung segmentation for interstitial lung disease using high resolution computed tomography (pp. 187–191).
21. Ren, Y. H., Sun, X. W., & Nie, S. D. (2010). A 3D segmentation method of lung parenchyma based on CT image sequences. In *Proceeding of 2010 International Conference on Information, Networking and Automation ICINA* (Vol. 2, pp. V2-332–V2-336).
22. S. P. Sahu, N. D. Londhe, and S. Verma, An Automated System for the Detection of Lung Cancer in CT data at Early Stages: Review.
23. Chen, K., Li, B., Tian, L., Zhu, W., & Bao, Y. (2014). Vessel attachment nodule segmentation using integrated active contour model based on fuzzy speed function and shape-intensity joint Bhattacharya distance. *Signal Processing*, 103, 273–284. Oct.
24. Kuhnigk, J.-., et al. (2006). Morphological segmentation and partial volume analysis for volumetry of solid pulmonary lesions in thoracic CT scans. *IEEE Transactions on Medical Imaging*, 25(4), 417–434.
25. Mukhopadhyay, S. (2016). A segmentation framework of pulmonary nodules in lung CT images. 86–103.
26. Wang, L., Lin, H., Huang, X., Wang, B., & Chen, Y. (2015). A 3d segmentation and visualization scheme for solid and non-solid lung lesions based on gaussian filtering regularized level set. In *Proceeding of 2014 International Conference on 3D Vision Work, 3DV, 2014* (pp. 67–74).
27. Paraagios, N., & Deriche, R. (1999). Geodesic active contours for supervised texture segmentation. In *Proceedings of 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)* (Vol. 2, pp. 422–427).
28. Santos, A. M., De Carvalho Filho, A. O., Silva, A. C., De Paiva, A. C., Nunes, R. A., & Gattass, M. (2014). Automatic detection of small lung nodules in 3D CT data using gaussian mixture models, Tsallis entropy and SVM. *Engineering Applications of Artificial Intelligence*, 36, 27–39.
29. Nithila, E. E., & Kumar, S. S. (2016). Segmentation of lung nodule in CT data using active contour model and Fuzzy C-mean clustering. *Alexandria Engineering Journal*, 55(3), 2583–2588.
30. Hao, R., Qiang, Y., & Yan, X. (2018). Juxta-Vascular pulmonary nodule segmentation in PET-CT imaging based on an LBF active contour model with information entropy and joint vector. *Computational and Mathematical Methods in Medicine*, 2018.
31. Chen, C. J., & Wang, Y. W. (2011). A preoperative 3D computer-aided diagnosis system for lung tumor. In *Proceeding of 2011 5th International Conference on Genetic and Evolutionary Computing ICGEC 2011* (pp. 279–282).
32. Dehmshki, J., Amin, H., Valdivieso, M., & Ye, X. (2008). Segmentation of pulmonary nodules in thoracic CT scans: A region growing approach. *IEEE Transactions on Medical Imaging*, 27(4), 467–480.

33. Oseas, A., et al. (2017). 3D shape analysis to reduce false positives for lung nodule detection systems. *Medical and Biological Engineering and Computing*, 55(8), 1199–1213.
34. Zhu, Y., Tan, Y., Hua, Y., Wang, M., Zhang, G., & Zhang, J. (2010). Feature selection and performance evaluation of support vector machine (SVM)-based classifier for differentiating benign and malignant pulmonary nodules by computed tomography. *Journal of Digital Imaging*, 23(1), 51–65.
35. Diciotti, S., Lombardo, S., Falchini, M., Picozzi, G., & Mascalchi, M. (2011). Automated Segmentation Refinement of Small Lung Nodules in CT Scans by Local Shape Analysis. 58(12), 3418–3428.
36. Muhammad, S., Muhammad, N., & Arfan, S. (2018). Lung nodule detection and classification based on geometric fit in parametric form and deep learning. *Neural Computing and Applications*, 3456789.
37. John, J., & Mini, M. G. (2016). Multilevel thresholding based segmentation and feature extraction for pulmonary nodule detection. *Procedia Technology*, 24, 957–963.
38. Rendon-Gonzalez, E., & Ponomaryov, V. (2016). Automatic Lung nodule segmentation and classification in CT images based on SVM. In *Proceeding of 9th International Kharkiv Symposium on Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves, MSMW 2016* (pp. 1–4).
39. Zhou, T., Lu, H., Zhang, J., & Shi, H. (2016). Pulmonary nodule detection model based on SVM and CT image feature-level fusion with rough sets. *Biomed Research International*, 2016.
40. Schilham, A. M., van Ginneken, B., & Loog, M. (2006). A computer-aided diagnosis system for detection of lung nodules in chest radiographs with an evaluation on a public database. *Medical Image Analysis*, 10(2), 247–258.



# Wavelet Decomposition Based Authentication Scheme for Dental CBCT Images



Ashish Khatter, Nitya Reddy and Anita Thakur

**Abstract** Incorrect information or documentation of the patient details and a wrong direction provided by any personnel to the patient can largely influence the lab reports or any other confidential data. Once the medical image is verified as secure and safe by the authorized access, one can review the medical image and the patient information. This paper focuses on providing a helping tool to ensure higher accuracy as well as authentication of the medical data and the related information. It proposes an authentication scheme to address the issue of security and privacy preservation of medical details. This technique involves generating a secure identification pattern from the fusion of the patient's information with the features of cone-beam CT images obtained through wavelet analysis without any tampering with the medical details. The patient's information can be successfully restored at the authorized receiver's end. This authentication scheme can be helpful and implemented in telecare medical information systems.

**Keywords** Medical image security · Authentication scheme · CBCT image · Wavelet decomposition · Medical information

## 1 Introduction

### A. Related Work

With the advancement and growing prevalence of communication technologies, telecare system enables transmission of data, and aids in medical treatment without any physical presence of the patient at the diagnosis center [1]. However, the safe and

---

A. Khatter (✉) · N. Reddy · A. Thakur  
Department of Electronics and Communication Engineering, Amity University, Noida, India  
e-mail: [ashish.khatter088@gmail.com](mailto:ashish.khatter088@gmail.com)

N. Reddy  
e-mail: [nitya1912reddy@gmail.com](mailto:nitya1912reddy@gmail.com)

A. Thakur  
e-mail: [athakur@amity.edu](mailto:athakur@amity.edu)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_53](https://doi.org/10.1007/978-981-15-0694-9_53)

secure transmission of medical information needs to be ensured. Improper documentation of the patient details or any wrong direction provided by any personnel to the patient can largely influence the lab reports or any other confidential data [2]. Therefore, the security of medical data is an important task in telecare medical information systems and authentication of information is an integral segment of the system [3, 4]. Furthermore, authentication of medical images, as well as patient information, should be done properly as any error can result in misdiagnosis and become a threat to the patient's life [5].

Various protocols and algorithms have been developed for authentication and privacy preservation of the medical information [6]. The technique of watermarking consists of three fragments. In the first one, the image is divided into two parts with one containing the information and the other with non-relevant data [7]. The second type is based on reversing the watermark in which no trace of watermark is left on the image [8]. In the third category of watermarking, some data is lost [9]. In all these techniques, some amount of degradation of the information takes place.

Many other methods have been implemented for resolving the security problems in telemedicine industry [10]. Biometric based authentication techniques have also been developed but, it has drawbacks of efficiency and cost [11]. Various algorithms have been developed which gave good results. Today, the most practical application is two-factor authentication [12]. Although, these works are interesting and have good results there is still a requirement to create new algorithms to strengthen the security for the medical industry. This work introduces a technique for solving the security problem of dental medical images without any loss of medical information.

## **B. Contribution**

The contribution of the proposed work is to generate a secure code or pattern from dental CBCT image to avoid any misuse of the medical information. Cone-beam computed tomography provides the ability to visualize three dimensional vital structures inside the mouth, for example, root canals, sinuses, and nerves. The details of the patient are combined with the features obtained from the wavelet analysis of the medical image. A unique secure digital pattern is generated. The receiver can verify and obtain the medical details of the patient through authorized access. This technique does not tamper any medical detail of the patient or the image. It can be implemented as an authentication scheme in telecare medical information systems.

## **C. Organization of Paper**

The paper is organized as follows. Section 2 presents the proposed methodology for authentication and privacy preservation of medical details. Section 3 displays the experimental results. Section 4 concludes the work and discusses future work in the related field.



Fig. 1 Dental CBCT image

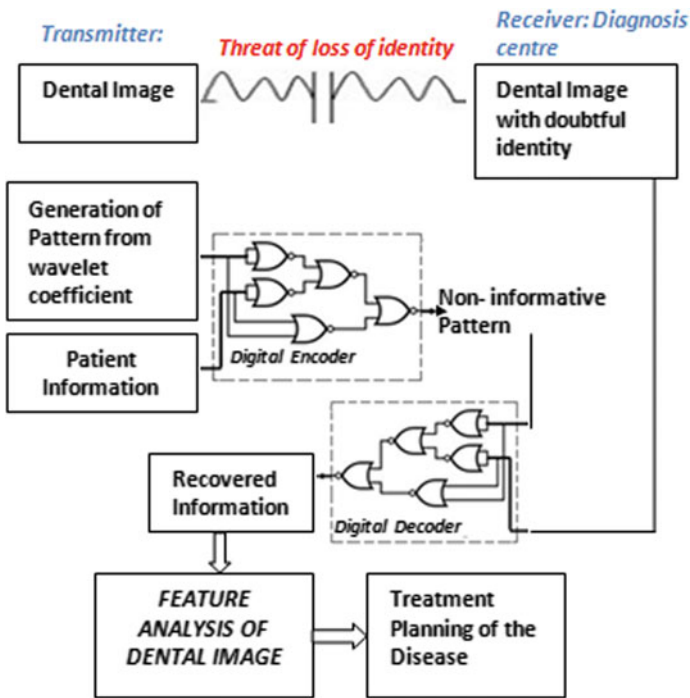


Fig. 2 Framework for the proposed methodology

## 2 The Proposed Methodology

The proposed methodology for authenticity and privacy preservation includes the wavelet analysis of the dental CBCT image. CBCT promises to revolutionize the procedures for the diagnosis and treatment of various dental ailments because of its technological advancement and diagnostic precision. It provides visualization

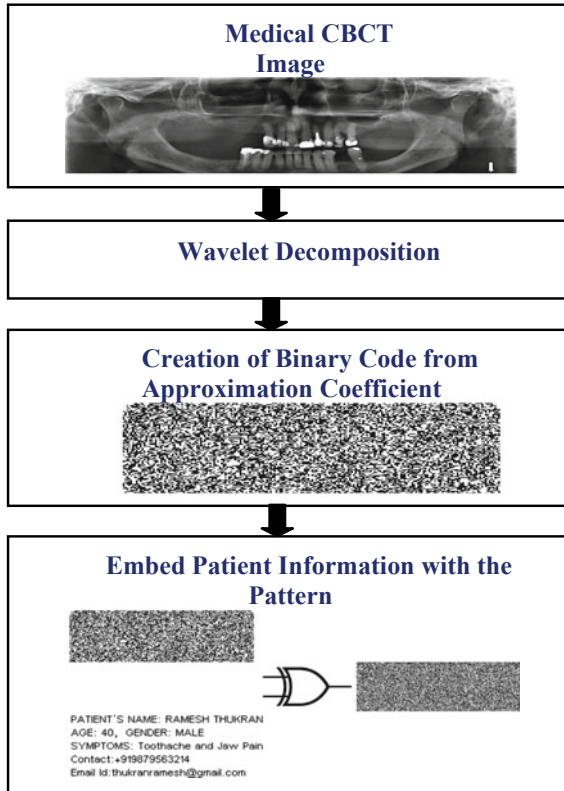


Fig. 3 Creation of the digital pattern from dental CBCT image

of three dimensional vital structures inside the mouth. For the patient, it aids in reduction of the treatment duration and offers better outcomes. Figure 1 shows a sample of dental cone beam computed tomography scan.

A binary pattern is created from the approximation coefficient of the decomposition vector obtained from the wavelet decomposition of the input medical image. This pattern is further embedded with the digital patient's information. As a result, a digital identification pattern is obtained without losing any important information. This pattern is non-informative which is sent to the receiver's center.

At the diagnosis center, the medical CBCT image is received firstly. The task is to recover the patient's details without any loss of medical information. The binary pattern is created from the dental image and then, the patient's information can be recovered with the help of the non-informative digital identification pattern (Figs 2, 3).

### 2.1 Algorithm: At the Transmitting Side

- Step1 Read the given dental CBCT image
- Step2 Divide it into  $4 \times 4$  blocks
- Step3 Perform wavelet analysis of each block (Multilevel wavelet decomposition)
- Step4 Obtain the approximation coefficient from the decomposition vector and calculate its mean value
- Step5 Create a binary pattern using the lowest significant bit of the mean
- Step6 Combine the digital patient's information with the binary pattern using exclusive OR function

Outcome **Digital Pattern Generated**

### 2.2 At the Diagnosis Center

- Step1 Read the received medical image
- Step2 Repeat Steps 2–5 as mentioned in the above algorithm

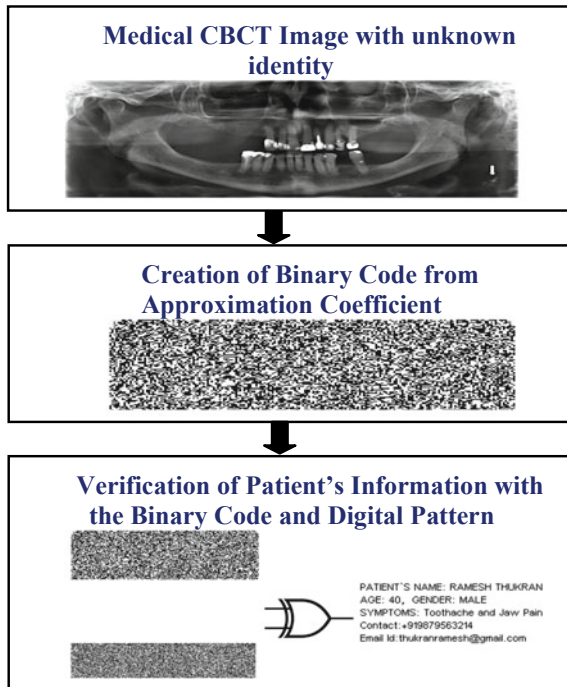


Fig. 4 Verification at the receiver side

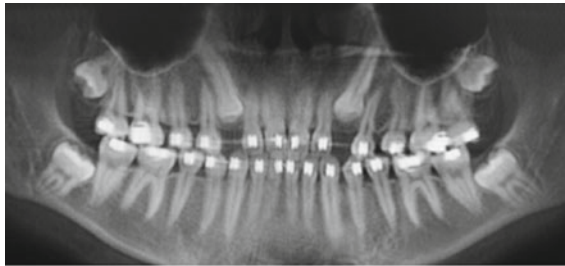
**Step3** Perform the logical “exclusive OR” operation with the binary code generated and the digital pattern obtained through authorized access.

**Outcome** **Patient’s Information Verified**

### 3 Results and Discussion

Experiments were performed on the publically available database of cone-beam CT scans. The database consists of 15 dental CBCT images [13].

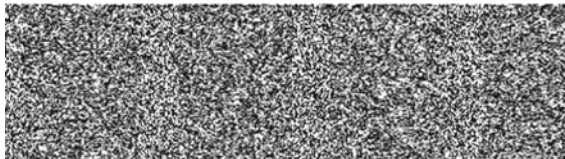
Following are the two samples of the CBCT image for identifying and diagnosing the disease, information of the patient, and the digital identification pattern obtained from them. Figures 4 and 5 display the patient medical information including the dental image as well the corresponding general details of two patients. The two details (medical image and details of the patient) are combined to form a unique digital pattern as shown in the figures.



(a)

PATIENT’S NAME: RAMESH THUKRAN  
 AGE: 40, GENDER: MALE  
 SYMPTOMS: Toothache and Jaw Pain  
 Contact: +919879563214  
 Email Id: thukranramesh@gmail.com

(b)



(c)

**Fig. 5** Sample I **a** Medical CBCT image **b** Information of the patient **c** Generated digital identification pattern from (a) and (b)

The recovery of patient details is done on the receiver center and the medical information of the patient can be verified successfully.

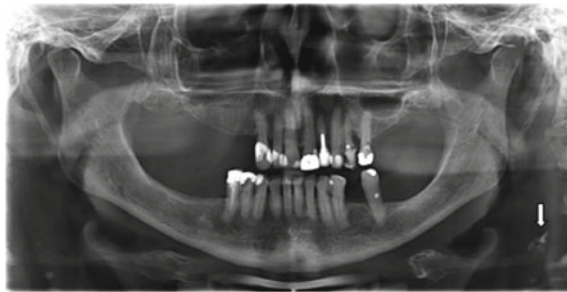
Figure 6 shows two restored digital information of the patients for the unknown dental image verification. It can be seen clearly that medical image is correctly recovered and there is no error in the information.. Henceforth, this algorithm can be implemented for image verification in telecare medical information systems.

**Uniqueness of Digital Identification Pattern:**

For the digital pattern of every medical image to be different and unique, the correlation coefficient has been calculated.

Table 1 shows the correlation coefficient of the digital pattern with the same pattern and with others. It has been calculated to find out that every medical image has a unique pattern.

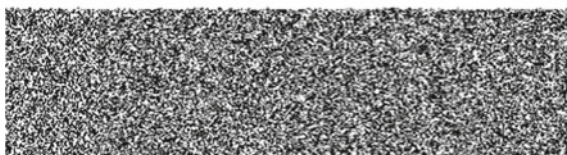
It is clear from the table that every digital pattern is unique. Figure 7 shows the bar graph for correlation coefficient for image 5. The correlation coefficient is 1 for the same pattern and approximately 0 for other samples (Fig. 8). Hence, a unique and different pattern is obtained for each and every individual medical image.



(a)

PATIENT'S NAME: ROSEATTE SEN  
AGE: 30, GENDER: FEMALE  
SYMPTOMS: ORAL PIERCING INFECTION  
Contact: +91 9741263214  
Email Id:redroseatte@hotmail.com

(b)



(c)

**Fig. 6** Sample II **a** Medical CBCT image **b** Information of the patient **c** Generated digital identification pattern from (a) and (b)

**Table 1** Uniqueness of digital pattern in terms of the correlation coefficient

Images	Correlation coefficient of digital pattern with the same pattern	Maximum correlation coefficient of digital pattern with different pattern
1	1.0000	0.0103
2	1.0000	0.0233
3	1.0000	0.0103
4	1.0000	0.0380
5	1.0000	0.0079
6	1.0000	0.0273
7	1.0000	0.0020
8	1.0000	0.0253
9	1.0000	0.0231
10	1.0000	0.0074
11	1.0000	0.0078
12	1.0000	0.0096
13	1.0000	0.0012
14	1.0000	0.0002
15	1.0000	0.0065

PATIENT'S NAME: RAMESH THUKRAN  
 AGE: 40, GENDER: MALE  
 SYMPTOMS: Toothache and Jaw Pain  
 Contact: +919879563214  
 Email Id: thukranramesh@gmail.com

(a)

PATIENT'S NAME: ROSEATTE SEN  
 AGE: 30, GENDER: FEMALE  
 SYMPTOMS: ORAL PIERCING INFECTION  
 Contact: +91 9741263214  
 Email Id: redroseatte@hotmail.com

(b)

**Fig. 7** Recovered data of the patients



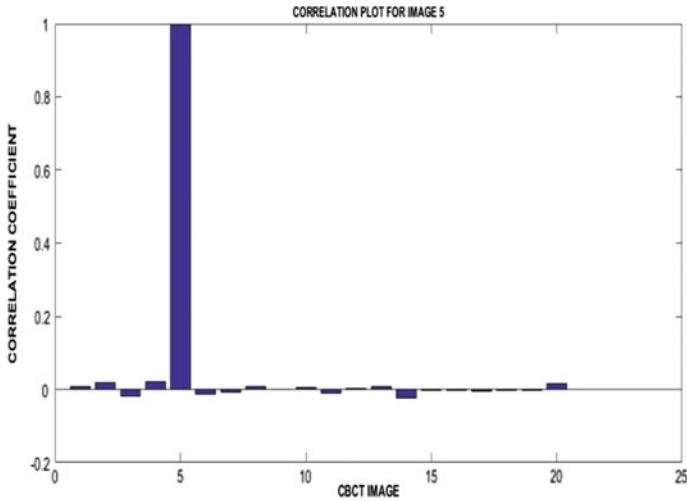


Fig. 8 Graph for correlation coefficient for image 5

## 4 Conclusion

The security of medical data is an important task in telecare medical information systems and authentication of information is an integral segment of the system. Furthermore, authentication of medical images, as well as patient information, should be done properly as any error can result in misdiagnosis and become a threat to the patient's life. The algorithm results in creation of a unique pattern from dental CBCT images through wavelet analysis function for authentication of the medical information. This technique involves generating secure identification pattern from the fusion of the patient's information with the features of cone-beam CT images obtained through wavelet decomposition without any tampering with the medical details. The patient's information is successfully restored at the authorized receiver's end. The computation and evaluation time of the pattern is less than a second. Hence, the method can be implemented as an authentication scheme in telecare medical information systems and the future work include research and experiment on other domains of the related fields.

## References

1. Tan, Z. (2018). Secure delegation-based authentication for telecare medicine information systems. *IEEE Access*, 6, 26091–26110. <https://doi.org/10.1109/ACCESS.2018.2832077>.
2. Xiong, H., Tao, J., & Yuan, C. (2017). Enabling telecare medical information systems with strong authentication and anonymity. *IEEE Access*, 5, 5648–5661. <https://doi.org/10.1109/ACCESS.2017.2678104>.
3. Li, C. T., Shih, D. H., & Wang, C. C. (2018). Cloud-assisted mutual authentication and privacy preservation protocol for telecare medical information systems. *Computer Methods Programs in Biomedicine*, 157, 191–203. <https://doi.org/10.1016/j.cmpb.2018.02.002>.
4. Achpazidis, I. (2008). *Image and medical data communication protocols for telemedicine and teleradiology*. Darmstadt, Germany: Department of Computer Science, Technical University of Darmstadt.
5. Al-Damegh, S. A. (2005). Emerging issues in medical imaging. *Indian Journal of Medical Ethics*, 2(4), <http://www.ijme.in/134co123.html>.
6. Kobayashi, L. O. M., & Furuie, S. S. (2009). Proposal for DICOM multiframe medical image integrity and authenticity. *Journal of Digital Imaging*, 22, 71–83. <https://doi.org/10.1007/s10278-008-9103-6>.
7. Tachakra, S., Wang, X. H., Istepanian, R. S. H., & Song, Y. H. (2003). Mobile e-health: The unwired evolution of telemedicine. *Telemedicine Journal and E-health*, 9(3), 247–257.
8. Memon, N. A., Chaudhry, A., Ahmad, M., & Keerio, Z. A. (2011). Hybrid watermarking of medical images for ROI authentication and recovery. *International Journal of Computer Mathematics*, 88, 2057–2071. <https://doi.org/10.1080/00207160.2010.543677>.
9. Tan, C., Ng, J., Xu, X., Poh, C., Guan, Y., & Sheah, K. (2011). Security protection of DICOM medical images using dual-layer reversible watermarking with tamper detection capability. *Journal of Digital Imaging*, 24, 528–540. <https://doi.org/10.1007/s10278-010-9295-4>.
10. Pan, W., Coatrieux, G., Cuppens-Boulahia, N., Cuppens, F., & Roux, C. (2010). Medical image integrity control combining digital signature and lossless watermarking. In J. Garcia-Alfaro, et al. (Eds.), *Data privacy management and autonomous spontaneous security* (pp. 153–162). Berlin: Springer.
11. Buciu, I., & Gacsadi, A. (2016). Biometrics systems and technologies: A survey. *International Journal of Computers Communications and Control* 11, 315–330. <https://doi.org/10.15837/ijccc.2016.3.2556>.
12. Grassi, P. A., Garcia, M. E., & Fenton, J. L. (2017). *Digital identity guidelines*. Gaithersburg, MD: National Institute of Standards and Technology.
13. Shah, N., Bansal, N., & Logani, A. (2014). Recent advances in imaging technologies in dentistry. *World Journal of Radiology*, 6(10), 794–807.

# A Comparative Analysis of Different Violence Detection Algorithms from Videos



Piyush Vashistha, Juginder Pal Singh and Mohd Aamir Khan

**Abstract** There are different methods or techniques used for identifying violence from video, such as hitting some object, kicking, fighting, and punching someone but still there is a big challenge for us to identify violence. However, some of the earlier mechanism generally extract descriptors around the spatiotemporal interesting points (STIP) or extract statistic features but there is limited effectiveness in detecting video-based violence. Therefore, the objective is to develop a better violence identification system that identifies the violence and triggers an alarm so that prompt assistance will be provided. This paper helps researchers who wish to study violent activity recognition and gather different insights on the main challenges and issues to solve in this emerging field.

**Keywords** Bag-of-words · Optical flow · Action recognition

## 1 Introduction

From the past few years, a lot of research has been done in the area of violence identification algorithms. The primary aim of analyzing video is to identify violence or unusual events with minimal or without human intervention.

Surveillance of video is an area of research that includes recognizing human events and their categorization into abnormal or normal activities.

There are three types of video surveillance systems:-

### 1. Manual

This surveillance system is totally based on person. This requires human labor for analyzing behavior to distinguish between normal and abnormal behavior.

### 2. Semi-automatic

Semi-automatic system needs minimal human interaction in comparison to manual surveillance system.

---

P. Vashistha (✉) · J. P. Singh · M. A. Khan  
GLA University, Mathura, India  
e-mail: [piyush.vashistha@gla.ac.in](mailto:piyush.vashistha@gla.ac.in)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_54](https://doi.org/10.1007/978-981-15-0694-9_54)

### 3. Semi-automatic

The fully automatic system does not need any human interaction in making any decision.

**Need of Violence Detection System** Nowadays many industries, government, private, and public sector are incurring a large amount of money for protecting their offices, premises, departmental stores, houses, etc. The primary aim of violence identification system is to identify unpredicted behavior that is coming from different types of violence.

Any event is unusual if it behaves differently from what we expect. Some of the examples are: a person hits another person, kicks another person, etc. These types of events generally have disordered movement or sudden movements.

Monitoring or tracking the full video stream by the human is not feasible due to time-consuming and tedious job; therefore, a self detection of unusual events in real time is required to prevent these types of activities. One of the solutions for detecting and tracking of motion is the concept of optical flow. This concept is used for segmenting the object and tracking of motion. Actions of a person can be visualized by histogram sequence of magnitude and orientation of optical flow.

For the evaluation or analysis of different violence detection algorithm, different datasets are used. Few of the datasets that are used in different violence identification system are CMU dataset, UTI dataset, PEL dataset, BEHAVE, HOF dataset, WED dataset, Hockey Fight database, etc.

## 2 Related Work

There are different types of frameworks which are available for detecting unusual behavior in Surveillance videos without any human interaction. Different researchers used various methodologies for identifying violence from the video sequence.

Datta et al. [5] proposed a methodology where they used motion path information for calculating the orientation of hands and legs. They also used the factor ‘Jerk’ to find some of the conclusions based on motion patterns. They addressed the challenges of identifying violence from different video sequence such as kicking, hitting some person, etc. For detecting violence, they depend on orientation information of human hand, leg, and information of motion path.

They defined a vector known as AMV comprising of direction and magnitude of motion. Also, they defined “Jerk” as the first derivative of Acceleration Measure Vector with regard to time. The conclusions and results are presented from various datasets that involve different sort of unusual activities. They also found the action of ‘pointing of finger’ is the prestage of the actual fight happening because, in finger pointing, the orientation of arm is parallel to the ground (Fig. 1).

Chen et al. [4] proposed their work which computes the “histogram of orientation of optical flow” and “the histogram of magnitude of optical flow”. In the first



Fig. 1 Finger pointing detection [5]

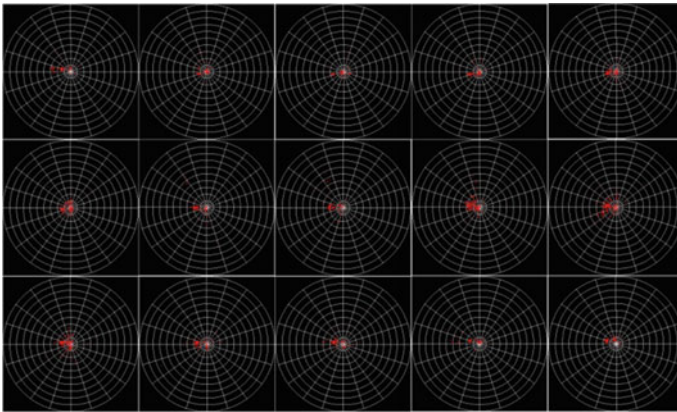


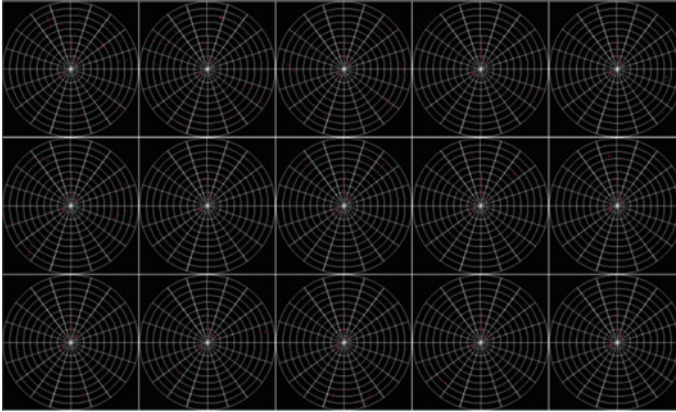
Fig. 2 Histogram of non-violence sequence [4]

Table 1 When machine learning method is random forest

		Predict	
		Fighting	Non-fighting
Observe	Fighting	49	1
	Non-fighting	3	47

step, from every frame, the features of optical flow are to be extracted. Then these characteristic points are dispersed in the entire optical flow system (Figs. 2 and 3).

They have calculated optical flows for each and every frame. A System of log-polar coordinate( $r$ ,  $\theta$ ) is used for estimating histogram. The radius ‘ $r$ ’ shows the magnitude of optical flow (0 to  $M_{max}$ ), where  $M_{max}$  is the maximum magnitude of optical flow. The angle ‘ $\theta$ ’ is the orientation of the optical flow (0 to 360). There are different Machine learning algorithms that are used in this research namely Bayesnet, SVM and Random Forest. There are 100 videos (50 fight videos and 50 non-fight videos) used for this research (Tables 1, 2, and 3). Their results are as follows:



**Fig. 3** Histogram of violence sequence [4]

**Table 2** When machine learning method is support vector machine

		Predict	
		Fighting	Non-fighting
Observe	Fighting	45	5
	Non-fighting	5	45

**Table 3** When machine learning method is support vector machine

		Predict	
		Fighting	Non-fighting
Observe	Fighting	42	8
	Non-fighting	6	44

Hassner et al. [10] proposed a system in which they used the concept of Violent Flow descriptor (ViF). The ViF descriptor is generated by computing the optical flow between each pairs of sequential frames. They calculated how the position of each pixel  $P_{x,y,t}$  changes from current frame (t) to next frame (t + 1). It also provides a flow vector  $(u_{x,y,t}, v_{x,y,t})$  of every pixel that matches with pixel in next frame. They considered only the magnitudes of the optical flow and it is calculated as:-

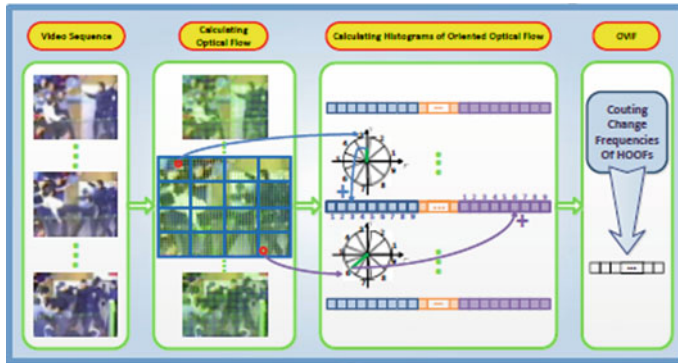
$$|m_{x,y,t}| = \sqrt{u_{x,y,t}^2 + v_{x,y,t}^2} \tag{1}$$

For every pixel of each frame, they obtained a binary indicator as

$$b_{x,y,t} = \begin{cases} 1 & \text{if } |m_{x,y,t} - m_{x,y,t-1}| \geq \theta \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

**Table 4** Results on crowd violence database

Method	Accuracy ( $\pm$ SD) (%)	AUC
LTP [12]	71.53 $\pm$ 0.17	79.86
HOG [13]	57.43 $\pm$ 0.37	61.82
HOF [13]	58.53 $\pm$ 0.32	57.60
HNF [13]	56.52 $\pm$ 0.33	59.94
ViF [2]	81.30 $\pm$ 0.21	85.00



**Fig. 4** Generating steps of OViF descriptor for video sequence [9]

Here,  $\Theta$  is a threshold for each frame which is the mean of  $|m_{x,y,t} - m_{x,y,t-1}|$ . It reflects the importance of change of magnitude between sequential frames. Then an average magnitude change map is obtained for every pixel for all the frames.

$$\bar{b}_{x,y} = \frac{\sum b_{x,y,t}}{T} \tag{3}$$

ViF [2] descriptor is the vector of quantized values  $\bar{b}_{x,y}$ . It is calculated by partitioning  $\bar{b}_{x,y}$  into  $m \times n$  cells and then assembling magnitude change frequencies in every cell separately. The change in magnitude in each cell is depicted by fixed size histogram and finally, these histograms are combined into a single vector (Table 4). Their experiment results are as follows:

Gao et al. [9] presented a new mechanism where they use the concept of ViF and OViF both. OViF represents the information which involves motion magnitude and motion orientation. The following figure shows the process of generating OViF descriptor (Fig. 4).

Firstly, optical flow is estimated between different pairs of consecutive frames of the given video sequence. Representation of every pixel of flow vector is as follows:-

$$|m_{x,y,t}| = \sqrt{u_{x,y,t}^2 + v_{x,y,t}^2} \tag{4}$$

**Table 5** Results on Hockey fight dataset [9]

Method	Classifier	Accuracy (SD)	AUC
LTP [12]	SVM	71.90 ± 0.49	–
ViF [2]	SVM	81.60 ± 0.22	88.01
	AdaBoost	73.70 ± 3.35	–
OVIF [9]	SVM	84.20 ± 3.33	90.32
	AdaBoost	78.30 ± 1.68	–
ViF+OVIF [9]	SVM	86.30 ± 1.57	91.93
	AdaBoost	82.30 ± 2.75	–
	AdaBoost+SVM	87.50 ± 1.70	92.81

**Table 6** Results on violent flows database [9]

Method	Classifier	Accuracy (SD)	AUC
LTP [12]	SVM	71.53 ± 0.17	79.86
ViF [2]	SVM	81.20 ± 1.79	88.04
	AdaBoost	77.60 ± 3.29	–
OVIF [9]	SVM	76.80 ± 3.90	80.47
	AdaBoost	74.00 ± 4.90	–
ViF+OVIF [9]	SVM	86.00 ± 1.41	91.82
	AdaBoost	82.40 ± 3.58	–
	AdaBoost+SVM	88.00 ± 2.45	94.84

$$\phi_{x,y,t} = \arctan \frac{v_{x,y,t}}{u_{x,y,t}} \quad (5)$$

The influence of ViF is limited in a special case like the flow vector of same pixel into two consecutive frames having the same magnitude but different directions. In such cases OVIF plays a significant role. This is because that ViF considers there is no difference between these two flow vectors, but actually they differ a lot (Tables 5 and 6).

Zhang et al. [17] proposed a classification model for improving the discrimination power of the dictionary. They used two terms namely “Representation Constraint” term and other is “Coefficient Incoherence” term. The “Representation Constraint” term is used to ensure a dictionary which has a good capacity when constructing a query image using various training samples which have an identical class level while the other term, i.e., “Coefficient Incoherence” term is used for ensuring a dictionary which has a poor capability which has different class labels. They have checked their results on the three datasets namely BEHAVE dataset, Crowd Violence dataset, and Hockey Fight dataset.



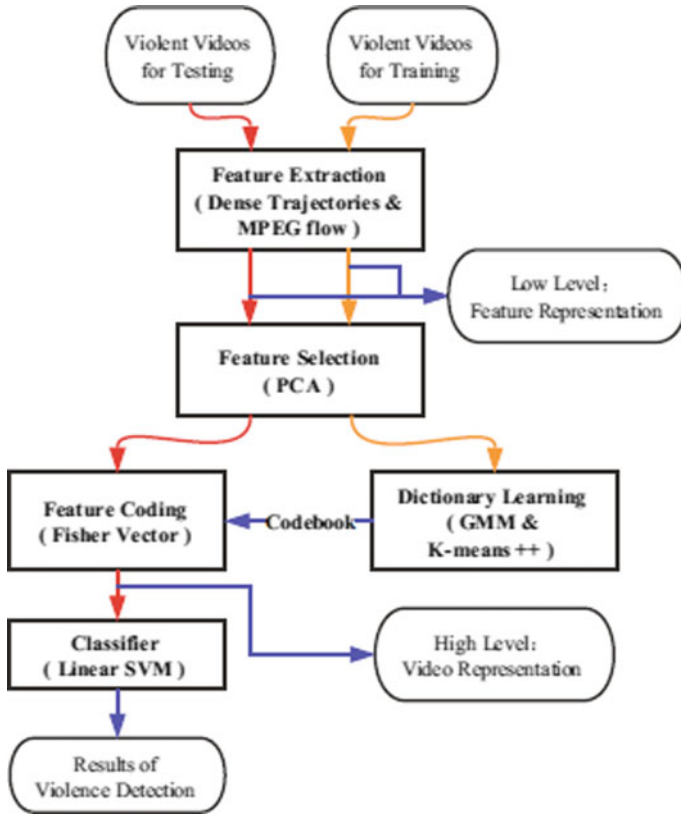


Fig. 5 Framework of violence detection [17]

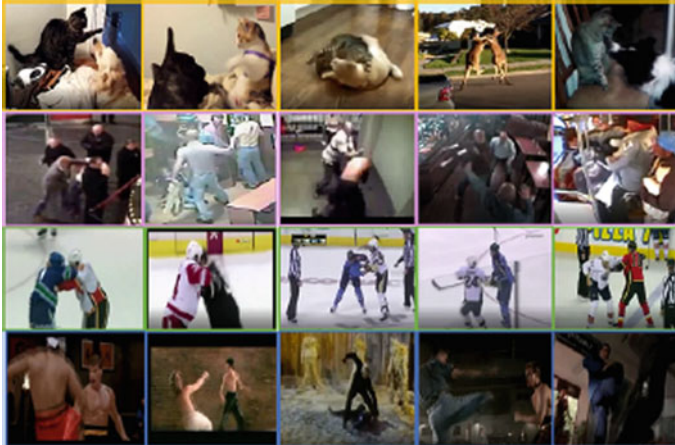
The performance of semi-supervised sparse classification approach is best for Hockey dataset. It results in a high recognition rate because of two terms namely “representation constraint” term and “coefficient incoherence term”.

In BEHAVE dataset, false alarms are generated when a crowd of people do violence or unusual activities. The results provided for this dataset is not satisfactory while the results for Crowd Violence dataset is robust and effective for identifying violence with composite scenarios like different distances from cameras, occlusions between crowd scenes, and people.

Cai et al. [3] proposed a system which uses Dense Trajectory and MPEG flow video descriptor algorithms. These algorithms are more robust and descriptive which describes appearance, shape, motion, and motion boundary. One another concept, i.e., Fisher Kernel is also used for transforming low-level features to high-level features. Fisher Vector is an effective algorithm for coding of the feature. This vector is derived from Fisher Kernel (Fig. 5).

**Table 7** Motion signals and statistics for features [7, 8]

Motion signals	Statistics description
Local motion magnitudes	Mean, maximum, minimum
Local motion accelerations	Median, standard deviation

**Fig. 6** Fight scenes from four fight datasets (animal fight, human fight, Hockey fight and action movies) [6]

They have used Principle Component Analysis, K-means, and Codebook size for improving the performance of video classification. The results on both the datasets namely Hockey dataset and Crowd Violence dataset is better than the other approaches. The accuracy in both the dataset is nearly 95% (Table 7).

Fu et al. [6] presented a cross-species learning method with a collection of low computational cost motion features for identification of violence or unusual activities. They proposed a set of useful local motion features (LMF), which includes motion statistics and segment correlation. The system extracts local motion features from each video. They used following different motion signals and statistics for features. They use ensemble classifiers to perform cross-species fight detection (Fig. 6).

They conducted experiments on 4 datasets: a public dataset containing human fights that occur in real life, a public dataset consists of human fights in hockey games, a public dataset capturing fight scenes from action movies and finally an animal fight dataset collected. These results are as follows (Table 8):

Vashistha et al. [15] proposed an approach where they used the combination of ViF and LBP [1]. The ViF vector computes the optical flow between each pair of sequential frames. It matches the position of each pixel of the current frame with the same pixel in the next frame. LBP helps in texture classification of the object. It helps in finding the local patterns of the moving object occurred in violence or unusual events. A  $3 \times 3$  neighborhood is used for a frame in LBP. It considers centered pixel

**Table 8** Evaluation of different feature representations for standard fight detection

Approach	Human fights (%)	Animal fights (%)	Hockey fights (%)	Action movies (%)
ViF [10]	80.1	80.1	81.6	93.0
OViF [9]	81.6	79.6	84.2	89.0
Motion signals [8]	82.7	79.6	84.5	98.5
LMF	89.8	85.0	87.5	99.0

for threshold and considered result in decimal number. Eventually, the texture of each is represented by a histogram of numbers ranging from 0 to 255.

Tay et al. [14] uses Convolution Neural Network to detect abnormal behavior. Their approach self learns the discriminative characteristics related to personal behavior from a huge collection of video sequences. In their proposed approach, they have taken RGB images from the video and undergone a preprocessing stage by applying a  $3 \times 3$  average filter for removing noise from the images. They have used CNN which has mainly three parts: the first part is the input layer that contains the input image, the second part is the middle layer which is used for detecting features, and the last part is the final layer which is acting as the classification layer.

They tested their results on different datasets namely CML Database, Web dataset and Hockey Fight dataset. Following Table 9 shows the different number of spans which are used to train with different learning rate parameter. This approach attains higher accuracy for all the above datasets. They have used the value of the learning rate parameter as 0.01 which gives the highest accuracy (Table 9).

Moreira et al. [11] proposed a multimodal fusion approach to sensitive scene localization in order to avoid the spread of inappropriate sensitive content on the internet like pornography and violence. The aim of this paper is to cease displaying inappropriate content and video. To validate their work, they performed localization experiments on pornographic and violent video content. They used the fusion of various and independent sensitive snippet classifiers. Each classifier relies on video data modality. The combination of snippet classifiers is carried out by fusing sensitive classification scores that are returned by each classifier.

Yang et al. [16] implemented and developed an aggression identification system which consists of mainly two parts: “a low-level observation part” and “a high-level reasoning part”. Experiments are done in a real train in cooperation with NS, the Dutch Railway Company to test the detection system. The prerequisite of their system was to detect unusual behavior patterns and the challenges were lighting conditions, train movement, background noise, and occlusions (Table 10).

**Table 9** Results obtained using different learning rate parameter

Learning rate	Maximum number of epochs	Dataset accuracy (%)				
		CMU	UTI	PEL	HOF	WED
0.001	10	99.66	54.90	90.38	58.5	89.21
	20	100.00	56.55	90.38	100.00	100.00
	30	100.00	57.74	90.38	100.00	100.00
	40	100.00	70.18	90.38	100.00	100.00
	50	100.00	99.15	90.38	100.00	100.00
	60	100.00	99.1	90.38	100.00	100.00
	70	100.00	99.7	90.38	100.00	100.00
	80	100.00	99.65	90.38	100.00	100.00
	90	100.00	99.7	90.38	100.00	100.00
	100	100.00	99.75	90.38	100.00	100.00
0.01	10	100.00	54.90	90.38	100.00	100.00
	20	100.00	99.6	90.38	100.00	100.00
	30	100.00	99.75	87.98	100.00	100.00
	40	100.00	99.55	100.00	100.00	100.00
	50	100.00	99.8	100.00	100.00	100.00
	60	100.00	99.8	100.00	100.00	100.00
	70	100.00	99.6	100.00	100.00	100.00
	80	100.00	99.7	100.00	100.00	100.00
	90	100.00	99.65	100.00	100.00	100.00
	100	100.00	99.8	100.00	100.00	100.00
0.1	10	46.64	54.90	9.62	50.00	0.00
	20	0.00	0.00	9.62	50.00	0.00
	30	0.00	54.90	0.00	50.00	0.00
	40	0.00	54.90	90.38	50.00	0.00
	50	0.00	0.00	9.62	50.00	50.00
	60	0.00	54.90	90.38	50.00	50.00
	70	0.00	0.00	9.62	50.00	50.00
	80	0.00	54.90	90.38	50.00	50.00
	90	0.00	54.90	90.38	50.00	0.00
	100	46.64	54.90	0.00	50.00	50.00
90	0.00	54.90	90.38	50.00	0.00	
100	46.64	54.90	0.00	50.00	50.00	

**Table 10** Parametric analysis of violence detection techniques

Paper	Method	Strength	Limitation
Person-on-Person violence detection in video data [5]	Acceleration measure vector and jerk	System has been carried out on a variety of people with different body built	System will break down if one of the person lie down
Fighting detection based on optical flow context histogram [4]	System of histogram of orientation and magnitude of optical flow	Higher accuracy achieved by Random forest classifier.	Not better when camera is moving
Violent flows: Real time detection of violent crowd behavior [10]	ViF	Accuracy is high	Fails when two contiguous frames have same magnitude
Violence detection using oriented violent flow [9]	ViF and OViF	Work even when two contiguous frames have same magnitude	The performance of OViF is not satisfying on violent flows database
Semi-supervised dictionary learning via local sparse constraints for violence detection [17]	Sparse classification model	Better for large-scale datasets	For BEHAVE dataset, results are not satisfactory
Violence detection based on spatiotemporal feature and fisher vector [3]	Dense trajectory, MPEG flow video descriptor and fisher vector	Better for both crowded scenes and noncrowded scenes	Dense trajectory process is time consuming
Cross-species learning: A low-cost approach to learning human fight from animal fight [6]	Local motion features	Differentiate violence across different species	To improve the performance of human fight detection, animal data is required
An Architecture to identify violence in video surveillance system using ViF and LBP [15]	ViF and LBP	Comparatively fast than previous approaches	Accuracy is low for Violent flow dataset
A Robust Abnormal Behavior detection method using CNN [14]	Convolution neural network	Accuracy is very high for all the datasets	Not good for crowd
Multimodal data fusion for sensitive scene localization [11]	Multimodal Data Fusion	Fusion pipeline approach is nicely adapted for each situation	This method misses about 5 min in every hour of streamed content
Automatic aggression detection inside trains [16]	Rule-based approach	Accuracy is high	Complexity in building and maintaining large rule-bases

### 3 Conclusion

This paper studies abnormal behavior detection in different situations. Most of the researchers used the concept of optical flow in their approaches. In all the approaches discussed above, the approach of CNN performs the best for abnormal behavior detection. The accuracy is nearly 95% for all the datasets while in previous approaches, accuracy is nearly 90%. CNN is a type of deep neural network used for analyzing the visual image. CNN automatically learns the characteristics which concern with the large range of abnormal behaviors. But this approach has a little drawback for crowd dataset. This paper discusses various issues related to the detection of violent activities. A summary of various different datasets that are used for violence detection is also listed in this paper. There are different models that were proposed to recognize various human activities related to violence. In each model, there were some limitations like some models does not work in complex situations like occlusion and crowded area. So in future, a robust surveillance system should be designed that can work for the crowd and different types of practical situations. So still a lot of research is ahead and continued improvements of such algorithms are also needed.

### References

1. Ahonen, T., Hadid, A., & Pietikäinen, M. (2004). Face recognition with local binary patterns. In *European Conference on Computer Vision* (pp. 469–481). Springer.
2. Arceda, V. M., Fabián, K.F., & Gutiérrez, J. C. (2016). Real time violence detection in video.
3. Cai, H., Jiang, H., Huang, X., Yang, J., & He, X. (2018). Violence detection based on spatio-temporal feature and fisher vector. In *Chinese Conference on Pattern Recognition and Computer Vision (PRCV)* (pp. 180–190). Springer.
4. Chen, Y., Zhang, L., Lin, B., Xu, Y., & Ren, X. (2011). Fighting detection based on optical flow context histogram. In *2011 Second International Conference on Innovations in Bio-inspired Computing and Applications* (pp. 95–98). IEEE.
5. Datta, A., Shah, M., & Lobo, N. D. V. (2002). Person-on-person violence detection in video data. In *Proceedings of 16th International Conference on Pattern Recognition* (Vol. 1, pp. 433–438). IEEE.
6. Fu, E. Y., Huang, M. X., Leong, H. V., & Ngai, G. (2018). Cross-species learning: A low-cost approach to learning human fight from animal fight. In *2018 ACM Multimedia Conference on Multimedia Conference* (pp. 320–327). ACM.
7. Fu, E. Y., Leong, H. V., Ngai, G., & Chan, S. (2015). Automatic fight detection based on motion analysis. In *2015 IEEE International Symposium on Multimedia (ISM)* (pp. 57–60). IEEE.
8. Fu, E. Y., Va Leong, H., Ngai, G., & Chan, S. (2016). Automatic fight detection in surveillance videos. In *Proceedings of the 14th International Conference on Advances in Mobile Computing and Multi Media* (pp. 225–234). ACM.
9. Gao, Y., Liu, H., Sun, X., Wang, C., & Liu, Y. (2016). Violence detection using oriented violent flows. *Image and Vision Computing*, 48, 37–41.
10. Hassner, T., Itcher, Y., & Kliper-Gross, O. (2012). Violent flows: Real-time detection of violent crowd behavior. In: *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (pp. 1–6). IEEE.
11. Moreira, D., Avila, S., Perez, M., Moraes, D., Testoni, V., Valle, E., et al. (2019). Multimodal data fusion for sensitive scene localization. *Information Fusion*, 45, 307–323.

12. Nanni, L., Brahnam, S., & Lumini, A. (2011). Local ternary patterns from three orthogonal planes for human action classification. *Expert Systems with Applications*, 38(5), 5125–5128.
13. Nievas, E. B., Suarez, O. D., García, G. B., & Sukthankar, R. (2011). Violence detection in video using computer vision techniques. In *International Conference on Computer Analysis of Images and Patterns* (pp. 332–339). Springer.
14. Tay, N. C., Connie, T., Ong, T. S., Goh, K. O. M., & Teh, P. S. (2019). A robust abnormal behavior detection method using convolutional neural network. In *Computational Science and Technology* (pp. 37–47). Springer.
15. Vashistha, P., Bhatnagar, C., & Khan, M. A. (2018). An architecture to identify violence in video surveillance system using vif and lbp. In *2018 4th International Conference on Recent Advances in Information Technology (RAIT)* (pp. 1–6). IEEE.
16. Yang, Z., & Rothkrantz, L. J. (2010). Automatic aggression detection inside trains. In *2010 IEEE International Conference on Systems, Man and Cybernetics* (pp. 2364–2372). IEEE.
17. Zhang, T., Jia, W., Gong, C., Sun, J., & Song, X. (2018). Semi-supervised dictionary learning via local sparse constraints for violence detection. *Pattern Recognition Letters*, 107, 98–104.

# Minimizing Synchronization Error in Compressed Domain Watermarking



Tanima Dutta, Aishwarya Soni and Hari Prabhat Gupta

**Abstract** The compressed domain video watermarking has been less expensive since complete decoding and then re-encoding is not required for watermark embedding. Handling synchronization error due to embedding is a challenging task, especially, due to motion compensation. In this paper, we propose an embedding scheme that can minimize desynchronization in the watermark extraction process at the decoder for embedding in compressed videos with better visual quality. Experimental results show the effectiveness of the proposed scheme.

**Keywords** Compressed domain embedding · Video watermarking · Synchronization error · Medical imaging · H.264/AVC videos

## 1 Introduction

In today's scenario, security is a major concern in various multimedia applications. To deal with such security issues, the watermarking technique is used which helps to hide the secret signature into digital videos. In Video Watermarking, embedded information should not be detected so that it has to be robust. Videos are, mainly, stored and transmitted in compressed format. It, therefore, increases the significance of compressed domain watermarking. Compressed video requires less computation because it does not need complete encoding and decoding in the embedding of watermark information. In recent years, various emerging technologies of video compression have been introduced [3]. Such technology aims to provide more compressed

---

T. Dutta · A. Soni (✉) · H. P. Gupta  
Department of Computer Science and Engineering,  
IIT (BHU) Varanasi, Varanasi, India  
e-mail: [aishwaryasoni.cse@gmail.com](mailto:aishwaryasoni.cse@gmail.com)

T. Dutta  
e-mail: [tanima.cse@iitbhu.ac.in](mailto:tanima.cse@iitbhu.ac.in)

H. P. Gupta  
e-mail: [hariprabhat.cse@iitbhu.ac.in](mailto:hariprabhat.cse@iitbhu.ac.in)



information as well as the improved visual quality of the compressed video. One of the compression technique is H.264/AVC which used in widespread applications [2, 4–7].

Synchronization error is a major problem in compressed domain video watermarking techniques which need to resist. Such error increases with more compressed videos like H.264/AVC. The parameters used in the embedding scheme of compressed video like motion vector, intra, and inter prediction modes values and quantized coefficients. Whenever re-encoding take place during the embedding process prediction mode value has been changed. This change affects the synchronization in watermark extraction with similar vales of parameters [9].

Synchronization error in the embedding process can reduce the robustness, therefore, it becomes simple to destroy watermark by attackers. This can be reduced only small visual quality. Additionally, synchronization error may create authentication issues. Location map is generally used to minimize synchronization error [8–11]. A location map is a file which contains the exact location of blocks or coefficients, where the watermark is embedded. Secured transmission of location map is still a potential overhead. The watermark can be completely destroyed if the attacker can able to access the location map.

There are various watermarking techniques that has been introduced in the last few years on H.264/AVC standard. Watermark techniques can be divided into two types; blind watermarking techniques [9] and non-blind watermarking technique [11]. When the watermark detection process requires original video for extraction it termed as *non-blind* technique. In a *blind* technique, the original video is not required for extraction of the watermark.

A non-blind watermarking scheme [11] introduced by Noorkami and Mersereau uses Watson visual model for embedding a watermark in I-frames. It is also used highly computation prediction process for watermark embedding and location map for finding a watermark coefficient or block. This technique is improved [10] for P-frames. It effectively improved the visual quality and also embedding take place in remaining non zero quantized P-frames. It used both the location map as well as original video to minimize the synchronization error.

Mansouri *et al.* introduced a blind watermarking scheme in H.264/AVC videos [9]. It used embedding of watermark in I-frames. Since blind watermarking methods are more preferable comparatively non- blind methods, it required the secure transmission of the original video, that increased overhead. In the blind scheme of watermarking, the selection of block is accomplished on the basis of luminance intra prediction mode value and number of nonzero coefficient of  $4 \times 4$  block. The watermark embedding method converts the nonzero coefficient to zero value as per hiding the watermark.

MINIMIZING SYNCHRONIZATION ERROR IN I- FRAMES: The intra prediction mode value altered [9] after re-encoding has performed because of synchronization error. The author [9] find the rate of intra prediction mode value altered from  $4 \times 4$  to  $16 \times 16$  for blocks having different number of nonzero residuals. When mode value change the block having nonzero residuals of large value is less in  $16 \times 16$ . Desynchronization has reduced when embedding is performed in macroblocks with

nonzero residuals of higher value. Further, other research authors described that the rate of intra mode change has reduced for  $4 \times 4$  of nonzero residuals after re-encoding.

**WHY P- FRAME EMBEDDING?** In the embedding method, P-frame provide effective visual quality as compared to I-frames [9, 11]. But I-frames have less risk of attack but the error during embedding has propagated throughout the uncoded blocks of I-frames due to intra prediction. It also occurs in P-frames and b-frames because of inter prediction in particular group of picture (GOP) block. hence it reduced the visual quality of the embedded video. The compressed video of H.264/AVC follows the GOP structure. Generally, it contains more numbers of B-frames than P-frames and has single I-frames. In B-frames, a maximum number of coefficient is zero because of bi-prediction motion estimation it becomes sparse. Embedding is performed with zero coefficient effect the video bit rate. Whenever the block is predicted from B-frames, the mode might be changed that cause enhancement in desynchronization. P-frames have less embedding capacity due to more number of P-frames having more numbers of nonzero coefficients but it has better visual capacity as compared to I-frames.

**SYNCHRONIZATION ERROR IN P- FRAME:** P-frames has both luminance intra and inter prediction modes. The inter prediction modes contain  $\{4 \times 4, 4 \times 8, 8 \times 4, 8 \times 8, 8 \times 16, 16 \times 8, \text{and } 16 \times 16, \text{SKIP}\}$  and luminance intra prediction modes having  $4 \times 4$  and  $16 \times 16$  block size. During embedding, ten different prediction mode values are changed so it could not be easily tracked.

**Our Contribution:** This paper is mainly focused on reducing the synchronization error in P-frames and enhances the visual quality. The compressed domain watermarking in P-frames require less computation as a result of not complete encoding and re-encoding in compressed video. As mainly, P-frames in compressed domain embedding give better visual quality comparatively other frames like I-frames [10]. The rate of alteration in prediction mode of P-frames has 10 different values, therefore, it is not easy to capture the changing of mode value after re-encoding that would decrease the synchronization. The block selection plays an important role in reducing the desynchronization during extraction process, therefore, compressed domain parameters like luminance prediction modes, number of nonzero coefficients, and motion vectors are carefully analyzed. For different video sequence, this experiment analyzes the result with and without using the location map.

The rest of the paper is organized as follows: Sect. 2.1 described the suitable block selection in embedding to minimize the synchronization error. Sections 2.2 and 2.3 presented the embedding and extraction scheme. Section 3 given the simulation results and conclusion of the paper in Sect. 4.

## 2 Proposed Scheme

The most challenging concern in P-frames is to minimizing the synchronization error in compressed domain embedding. At the decoder side during the watermark extraction process, prediction error has propagated in the remaining uncoded blocks of

P-frames. The suitable block should be selected for embedding which can minimize the desynchronization. The size of the block is  $4 \times 4$  in the rest of this paper.

## 2.1 Block Selection

In Sect. 1 describe that P-frames have 10 different prediction modes. For selecting suitable blocks there exist different parameter to analyze such as luminance prediction modes, number of nonzero residuals, and motion vectors. Here, to reduce the desynchronization it is very essential to select the appropriate blocks. Block selection is performed on the following standard.

- **SKIP mode blocks:** During embedding in P-frames avoid zero coefficient block such as SKIP mode blocks. When the embedding is performed in zero coefficients it decreases the visual quality and increases the video bit rate. Also, desynchronization is increased while changing SKIP mode to another mode by changing the zero coefficient to nonzero coefficient block.

- **Intra  $16 \times 16$  and Inter  $16 \times 16$  blocks:** There is a high likelihood of intra  $16 \times 16$  and inter  $16 \times 16$  mode block which can be changed from nonzero coefficient to zero coefficient. Therefore, SKIP mode extraction cannot be done correctly. At decoder side, zero coefficient blocks should be avoided in embedding due to increase in deynchronization.

- **Number of Nonzero Quantized Residuals:** Similarly [9], selection of blocks for embedding scheme that can reduce the synchronize error will be done by the number of nonzeros quantized residuals. It means if there are more number of nonzero coefficient, there will be less synchronization error. In intra coded luminance blocks, more number of nonzero quantized residuals are present, whereas inter coded luminance blocks have less number of nonzero quantized residuals. The nonzero quantized residuals having larger number is less in  $4 \times 4$  of P-frames. When nonzero quantized residuals is larger in number then the number of blocks for embedding is decreased [9]. The spatial characteristics of the video sequence is used to select the suitable threshold [9]. Since the number of nonzero quantized residuals in the block of P-frames are more than eight which is very rare, so it cannot be examined for comparison.

- **Motion Vector Field:** Compressed domain watermarking in P-frames, error propagate in uncoded blocks. It is also necessary to avoid temporal flicker by exploiting the motion information. The low motion regions are more effective against synchronization error. Generally, motion regions are in encoded form. When low motion blocks change into high motion blocks cause to change the prediction mode after re-encoding. In general, P-frames of  $4 \times 4$  blocks having higher numbers of nonzero residuals in intra mode is suitable for embedding [9]. In Intra mode, predicted blocks of P-frames have zero motion vectors. It is necessary to estimate the motion vector field of P-frames block. To estimate complete motion vector field for all blocks in P-frames, first, normalize each motion vector to ensure that it points directly to the location in the previous frame [1]. In H.264/AVC standard have multiple

referencing frames. Therefore, normalization ensures the referencing relationship of frames. After normalization,  $3 \times 3$  median filter is used to smoothen motion vectors to get complete motion vector field.

## 2.2 Watermark Embedding

The watermark will be embedded with selecting appropriate blocks but it also, need to select some parameters like number of nonzero quantized residuals and sign and value of quantized residuals as well as motion vectors for the actual embedding of watermark. There are following parameter which may be avoided because they are against desynchronization:

- **Changing the number of nonzero quantized residuals:** For embedding watermark, it is necessary to change the nonzero coefficients to zero coefficients [9]. When nonzero coefficients changed to zero coefficients predicted blocks can change the mode which enhances the desynchronization. Additionally, It also reduces the visual quality. The number of nonzero residuals changes by changing the zero coefficients to nonzero coefficient hence, it again increased the synchronization error. This change also significantly increase in video bit rate.

- **Changing the value of motion vector:** For Fragile watermark, need to change motion vectors for embedding [8]. Therefore, the blocks which are predicted from such blocks may be distinct reference blocks. This decreases synchronization significantly.

- **Changing DC residuals or sign of quantized AC residuals:** This resultant predicted block from such blocks could change the modes, as it caused an increase in synchronization error and also highly degrades the visual quality [10].

In the above mentioned discussion, it is clear that a suitable block for watermark embedding scheme is changing the value of quantized AC residuals block and keeps the same sign of the coefficient [10]. Unlike [10, 11], Whenever the original video does not need for extraction process at decoder, then embedding scheme should be blind ([9]). In this present scheme, such  $4 \times 4$  blocks are used for embedding the watermark after selection (Sect. 2.1). During embedding first of all, select two nonzero quantized AC coefficients in candidate block represented by  $C_1$  and  $C_2$  for watermark embedding. Also,  $C_1$ ,  $C_2$ , and  $t$  is used in each P-frame blocks for watermark embedding as follows.

On the basis of visual quality, the value of threshold  $t$  is decided. In a smooth region, the value of threshold  $t$  is small while in high textured regions, it is higher. Absolute value of  $z$  is represented by  $|z|$ .

---

**Algorithm: Watermark Embedding**

---

```

if watermark bit is 0 then
  if  $|C_1| \leq |C_2|$  then
     $|C_1| = |C_2| + t$ 
    sign of the coefficient remain same
  end
else
  if  $|C_1| \geq |C_2|$  then
     $|C_2| = |C_1| + t$ 
    sign of the coefficient remain same
  end
end
end

```

---

### 2.3 Watermark Extraction

Watermark extraction method is an opposite process of watermark embedding. It is possible after entropy decoding at the decoder. The procedure through which watermark bit is extracted from each candidate blocks in watermark video is as follows:

$$\text{The watermark bit} = \begin{cases} 0 & |C_1| > |C_2| \\ 1 & \text{otherwise} \end{cases}$$

## 3 Experimental Results

In this paper, the implementation of the presented scheme used H.264 standard software [12] and also used {foreman, carphone, news, trevor} videos in evaluation. Frame resolution of Quarter Common Intermediate format (QCIF) is  $176 \times 144$  which is used in video sequences. In general, 30 frames per second as frame rate was used. The value of a parameter such as quantization Parameter (QP) is 28, intra period is 10, and group of picture (GOP) used {IBBPBBPBBP}. All prediction modes such as intra and inter modes are enabled and employed high profile as well as fast full search in JM. The minimum value such as the number of nonzero quantized residuals and nonzero quantized residuals are 3 and 2 whereas the minimum range is considered as 3 to 15 of motion vector field.

On the basis of Peak Signal-to-Noise Ratio (PSNR), visual quality of the watermarked video compared with other existing scheme. In Fig. 1, it represents the average Peak Signal-to-Noise Ratio used payload = {100, 150, 200, 250, 300} of average 100 frames per video, for the proposed scheme with existing scheme. From the result (Fig. 1), it is clear that P-frames in the embedding scheme gives better visual quality as compared to I-frames.

Evaluation of robustness against different attacks of proposed scheme is based on bit error rate.

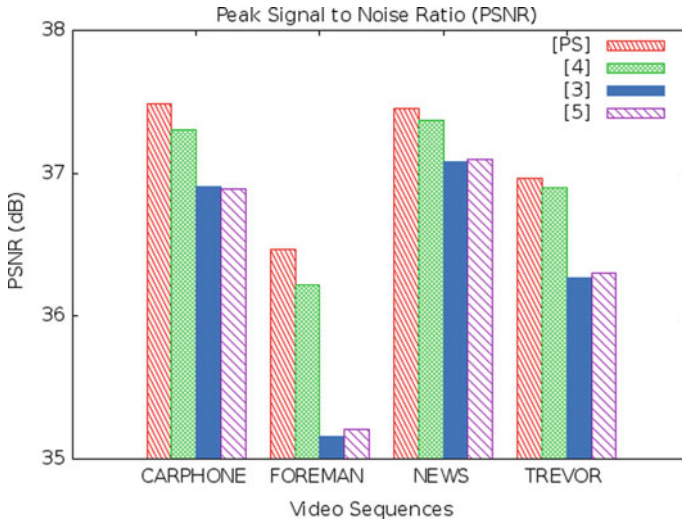


Fig. 1 Peak Signal-to-Noise Ratio (PSNR)

**Bit Error Rate (BER):** Bit Error Rate is defined as the frequency of bit errors when detecting a multibit watermark message [13]

$$BER = \frac{\text{number of error bits}}{\text{total number of bits sent}} \tag{1}$$

Robustness of a watermarking method [13] is given by

$$\text{Robustness} = (1 - \text{BER}) \times 100. \tag{2}$$

The palette where the information of embedding location is saved [8–11]. The palette is needed in the extraction of watermark in decoding. Without palette, watermark location is selected as incorrect, therefore, there will be more chance of loss in synchronization in watermark sequence and make less robustness of watermark method.

In Fig. 2, it shows the comparison of robustness results in blind scheme (proposed scheme [PS], [9]) between recompression error with and without providing a palette to the decoder. The robustness in case of recompression error is acceptable in the proposed scheme. Figures 3 and 4 shows the comparative results of robustness with and without providing a palette to the decoder by “changing quantization parameter (QP)” from 28 to 26 and 30, respectively. The proposed method help in minimizing desynchronization similar [9], in P-frames by changing the quantization parameter.

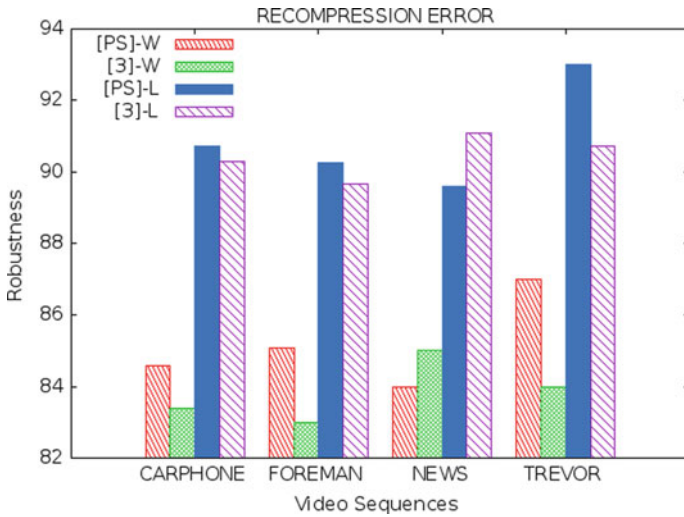


Fig. 2 Robustness against recompression error with (-L) and without (-W) using palette

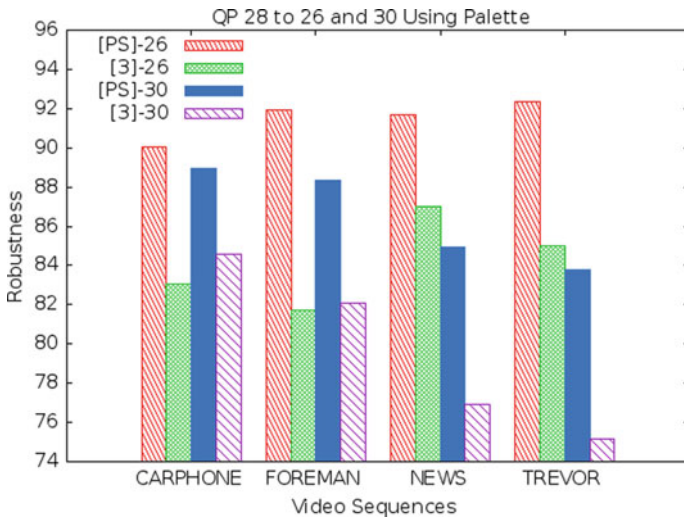


Fig. 3 Robustness against changing QP 28 to 26 and QP 28 to 30 using palette

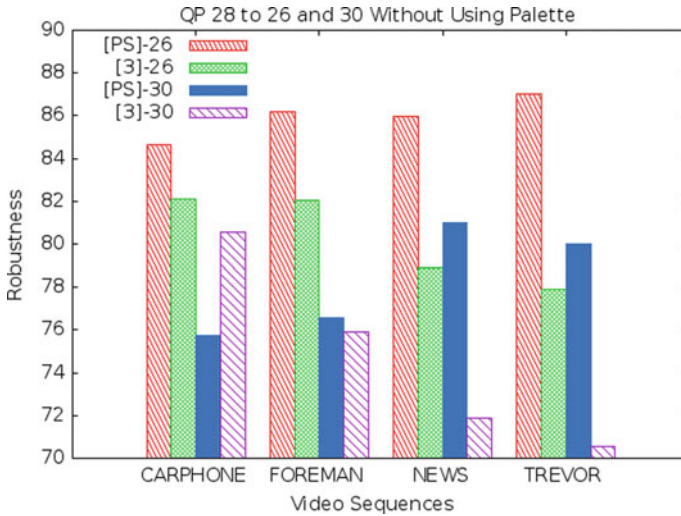


Fig. 4 Robustness against changing QP 28 to 26 and QP 28 to 30 without using palette

### 4 Conclusion

In this research paper, desynchronization is minimized in compressed domain watermarking at the decoder side during the extraction of watermark in P-frames. The proposed scheme is effective in minimizing synchronization error and also relating to robustness and visual quality which can be seen in Simulation results.

In further, the robustness of the proposed method can be identified by applying with different watermarking attacks like collusion and copy attacks and similar image/video processing attacks. The state of art, literature is compared with several concerns such as security and enhancement in video bit rate.

**Acknowledgements** This work is supported by Science and Engineering Research Board (SERB) file number ECR/2017/002419, project entitled as A Robust Medical Image Forensics System for Smart Healthcare, and scheme Early Career Research Award.

### References

1. Dong, P., Xia, Y., Zhuo, L., & Feng, D. (2011). Real-time moving object segmentation and tracking for H.264/AVC surveillance videos. In *Proceedings of ICIP*, pp. 2309–2312.
2. Dutta, T. (2013). Motion compensated compressed domain watermarking. In *Proceedings of ACM International Conference on Multimedia*, pp. 1039–1042.
3. Dutta, T. (2015). Medical data compression and transmission in wireless Ad Hoc networks. *Sensors Journal, IEEE, 15*, 778–786.



4. Dutta, T., & Gupta, H. P. (2016). A robust watermarking framework for High Efficiency Video Coding (HEVC) Encoded video with blind extraction process. *Journal of Visual Communication and Image Representation*, 38, 29–44.
5. Dutta, T., & Gupta, H. P. (2017). An efficient framework for compressed domain watermarking in P frames of high-efficiency video coding (HEVC)–encoded video. *ACM Trans Multimedia Computer and Communications Application*, 13(1), 12:1–12:24.
6. Dutta, T., Sur, A., & Nandi, S. (2013a). A robust compressed domain video watermarking in P-frames with controlled bit rate increase. In *Proceedings of National Conference on Communications (NCC)*, pp. 1–5.
7. Dutta, T., Sur, A., & Nandi, S. (2013b) MCRD: Motion coherent region detection in H.264 compressed video. In *Proceedings of IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6.
8. Feng, G., & Wu, G. (2011). Motion vector and mode selection based fragile video watermarking algorithm. In *Proceedings of ASID*, pp. 73–76.
9. Mansouri, A., Aznavah, A., Torkamani, F., & Kurugollu, F. (2010) A low complexity video watermarking in H.264 compressed domain. *IEEE Trans on Information Forensics and Security*, 5(4), 649–657.
10. Noorkami, M., & Mersereau, R. (2008). Digital video watermarking in P-Frames with controlled video bit-rate increase. *IEEE Trans on Information Forensics and Security*, 3(3), 441–455.
11. Noorkami, M., & Mersereau, R. M. (2007). A framework for robust watermarking of H.264-encoded video with controllable detection performance. *IEEE Trans on Information Forensics and Security*, 2(1), 14–23.
12. Sahrng, K. (2019). H.264 Reference Software Group. <http://iphome.hhi.de/suehring/tml/>.
13. Salomon, D., Motta, G., & Bryant, D. (2010). An Engineer’s guide to Automated Testing of High-speed Interfaces. Artech House.

# Deep Learning Architectures for Computer Vision Applications: A Study



Randheer Bagi, Tanima Dutta and Hari Prabhat Gupta

**Abstract** Deep learning has become one of the most preferred solution for many complex problems. It shows outstanding performance in the field of computer vision to perform tasks like, image classification, object detection, and image generation. Recently, many research efforts are focused on changing the deep learning architecture for widespread application domain. In this paper, we present a comprehensive survey on the various issues and challenges faced by deep learning techniques. Furthermore, we analyze different deep learning architectures to provide the solution for the computer vision tasks along with their importance.

**Keywords** Deep Learning Architecture · Computer Vision · Convolutional Neural Network (CNN) · Recurrent Neural Network (RNN)

## 1 Introduction

With the advancement in technology the utility of digital appliances is increasing in our day to day life. This advancement, draws up the focus of researchers and industry personals. The digital world has a lot of data in the form of image, video, text, and audio. Machine learning is introduced to handle this large digital world data and provide an appropriate solution. Traditionally, machine learning relies highly on data preprocessing and its representation techniques. Therefore, it requires domain expertise for data preprocessing, feature extraction, and feature selection. The hand crafted feature generation has a limited scope because features vary according to different applications. Before the emergence of deep learning, feature engineering

---

R. Bagi (✉) · T. Dutta · H. P. Gupta  
Department of Computer Science and Engineering, IIT (BHU) Varanasi, Varanasi, India  
e-mail: [randheerbagi.rs.cse17@iitbhu.ac.in](mailto:randheerbagi.rs.cse17@iitbhu.ac.in)

T. Dutta  
e-mail: [tanima.cse@iitbhu.ac.in](mailto:tanima.cse@iitbhu.ac.in)

H. P. Gupta  
e-mail: [hariprabhat.cse@iitbhu.ac.in](mailto:hariprabhat.cse@iitbhu.ac.in)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_56](https://doi.org/10.1007/978-981-15-0694-9_56)

using machine learning was at its boom [10, 33]. The feature engineering however required more human effort and also specific domain knowledge.

Deep learning is a subset of machine learning algorithms. The growth of deep learning has been substantial because of its ability to deal with huge data size and has also produce a significantly improved result. In recent years, deep learning has become more popular due to its wide applications like computer vision, multimedia security [11], information retrieval, image classification, medical imaging, and scene text detection. Deep learning is inspired by neurons of the human brain [6]. The main motive behind developing deep learning is to mimic the functionality of a human brain. The architecture in deep learning mainly uses graph representation, i.e., multiple nodes and neurons are in network connected to each other at multiple layers to learn different models. Deep network improves the feature generation procedure. It extracts high-level and low-level features. This automatic generation of features minimizes the error which improves the accuracy.

The main contributions of this paper are summarized as follows:

1. We review the deep learning techniques and its emergence over classical machine learning techniques. We also discuss the functionality of deep learning.
2. We explain the several computer vision applications. We also discuss the different deep learning models which incorporated the computer vision applications.
3. We provide a complete insight into the various architectures of deep learning techniques. We demonstrate the deep learning models in detail along with their advantages and limitations.

The rest of the paper is organized as follows: first, we present a study of deep learning architecture, applications, issues and challenges. Next, we focus on area of computer vision in which deep learning is used for object detection, image classification, and image reconstruction. In the Sect. 2, we investigate the basics of computer vision tasks and deep learning operations. The architectural modifications and work-flow of various deep learning models are depicted in Sect. 3. Different architectures and datasets are discussed in Sect. 4. We conclude the paper in Sect. 5.

## 2 Preliminaries

In this section, we first describe the applications of computer vision which is followed by the utility of deep learning in the field of computer vision. Computer vision is defined as providing the capability to the machine that replicates the human vision system [4, 6]. Some of the computer vision tasks are as follows:

1. **Object classification:** In the object classification, we add a label to the input image according to their class score.
2. **Object localization:** After the classification, if the classified image has only one object within the image, we need to localize the object. The task of finding the location of the object in the given image is called localization.

3. **Object detection:** It is a task of identifying multiple objects within the given image. The object may belong to the same group or another class group.
4. **Image segmentation:** It is defined as the grouping of pixels that have similar properties. The segmentation algorithm mostly depends on its applications.

Deep networks having large number of parameters are used to perform the computer vision task. Some of the important deep network operations used in computer vision are described as follows:

1. **Convolution (conv):** Convolution operation is performed on an input image for the computation of feature maps. Fixed size of kernel slides through the complete input image. The kernel is a weighted random matrix which is updated in stochastic gradient (SGD) process for parameter tuning to minimize the error in accuracy [9, 19, 28].
2. **Max-pooling:** Max-pooling is performed in the deep network to minimize the network complexity by selecting the relevant features from the feature map. Such selection causes the loss of low level feature information in deep network [9, 19, 28].
3. **Padding:** Padding is used to handle the error introduced in the conv operation i.e., output shrink and information loss from the corners of the image. It takes care of dimensionality mismatch of data between layers of deep network [9, 19, 28].
4. **Dropout:** It is introduced to prevent the problem of over-fitting in deep network. It is a regularization technique. In the testing phase, we do not use dropout algorithm in the network [9, 19, 28].
5. **Back-propagation:** It is used to minimize the cost function of deep network. It is known as back-propagation because its computation starts from the output layer and ends at the input layer. In this, we modify the weights, biases, or activation function between the layers of deep network to reduce the error in accuracy [9, 19, 28].
6. **Activation function:** It is used to add non-linearity in the deep network. By adding the non-linearity it solves the problem of vanishing gradient. In vanishing gradient problem we do not find the suitable parameter to update the network during back-propagation. Widely used activation functions are relu, tanh, elu, sigmoid etc [9, 19, 28].

### 3 Deep Network Architecture

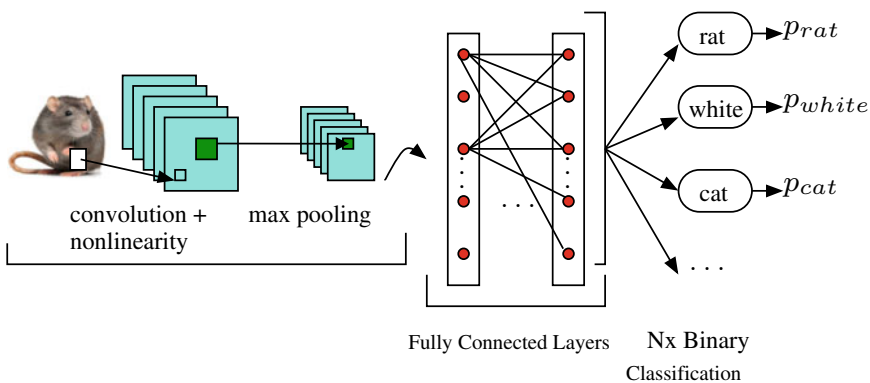
In this section, we describe various deep network architectures which are widely used in the field of computer vision with their importance. At the first layer of deep network architecture we take the input. The middle layers are also known as hidden layers. As the number of hidden layers increase, the architectures transform from shallow to dense network. The output layer has a fully connected layer and a softmax activation function.

### 3.1 Convolutional Neural Network (CNN)

It is a kind of multi-layer neural network, designed to recognize visual pattern directly from pixel of images with minimal preprocessing. The architecture takes input as an image and produces an output as a class score. It has a sequential structure of conv, Relu, and maxpool [20, 21]. Finally, at the end of the network it has a fully connected layer with softmax activation function as shown in Fig. 1. In the fully connected layer all the layers are stacked in a one dimensional array. The softmax is used to estimate the score for the prediction of class level of input test image. It is the most stable architecture of deep learning model. Many other versions of deep network architecture are developed by manipulating the CNN architecture. It uses  $3 \times 3$  size filter to perform the convolution in convolution layer. Maximum size of receptive field in maxpooling is restricted to  $3 \times 3$  because when we further increase receptive field size it causes the loss of information from the input data [15].

### 3.2 AlexNet

AlexNet is one of the basic deep learning model. It is used in computer vision for the tasks like classification, localization, and detection [19]. It has five conv layers and max-pooling layers in a successive way and then a fully connected layer at the end of the network. Relu activation function is applied after every convolution layer to add non-linearity in the network of AlexNet. Data augmentation techniques are used to avoid biased input data. It performs image translations, horizontal reflections, and patch extractions. Two GPUs are used in parallel to train the model and the parameters are tuned by using SGD process.



**Fig. 1** Working of convolution neural network [21]

### 3.3 ZF Net

ZF Net comes with a more refined architecture of the AlexNet. It minimizes the loss of important information at the initial layer by reducing the filter size from  $11 \times 11$  to  $7 \times 7$ . Along with conv layers, it also has deconvolutional layers which are used to visualize the feature map [35]. Non linearity is added in the network using relu activation function. Here, cross entropy is used to measure the error of the loss function. It uses SGD to tune the parameters for minimizing the error in accuracy.

### 3.4 VGG Net

It was introduced with the idea to minimize the number of parameters in the network of ZF Net architecture. It therefore uses a filter size of  $3 \times 3$ . The use of small filter size increase the depth of the network and reduces the dimension of input data at each layer [29]. In this network the filters are doubled after each maxpool layer. It refines the result of image classification and localization. Here, batch gradient descent is used for parameter tuning in training phase of the network. Relu activation function is used to add non-linearity. The model is trained over four GPUs.

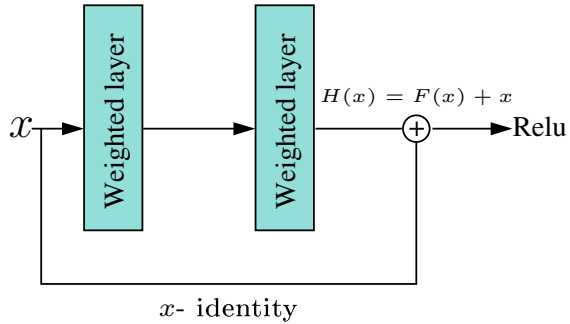
### 3.5 GoogLeNet

It is an extension of the CNN model. It focuses on CNN architecture and comes out with the conclusion that the stacking of multiple layers causes computational and space complexity [31, 32]. Therefore, in this network some of the sequential conv and max-pooling operations are converted to parallel operations. In this architecture,  $1 \times 1$  conv operation is also added to every conv and max-pooling operation. A block of parallel operations with  $1 \times 1$  conv operation is called an inception module. The inception module replaces the max pool operation with average pooling. This operation minimizes the computational parameters in the network. The class score is predicted by the softmax function in the network.

### 3.6 Microsoft ResNet

ResNet is introduced with an observation that when we keep adding layers in the network of the deep architecture, the accuracy of the model decreases [12, 17]. Therefore, to counter this issue in the network, after every two layers a skip connection is added. This skip connection with two layers is known as a residual block as shown in Fig. 2. Here, each block consists of a series of layers with a shortcut connection. It provides more control over the network, because in an ultra-deep

**Fig. 2** Residual block of ResNet [12]



network without using skip connection it is difficult to choose the tuning parameter. It decreases the error function exponentially because of the use of skip connection in back-propagation. It uses eight GPUs for the training of the model.

### 3.7 Region Based CNN (RCNN)

This architecture is proposed for object detection in the given input image by selective search [24]. The basic model is known as RCNN, then Fast-RCNN and Faster-CNN is proposed and finally, Mask-RCNN. Mask-RCNN is simply adding a binary mask to the output of the Faster R-CNN. This binary mask ensures the detected pixel is a part of object or not. Significant accuracy is achieved in the state of art of object detection by using Faster-RCNN but it takes larger testing time. In the basic network of deep learning i.e., CNN, Faster-RCNN comes with Region Proposal Network (RPN). An RPN takes input as an image and generates a set of rectangular box, these boxes are well known as object proposal and each box has an objectness score. Region proposals are generated by sliding a small network over the convolutional feature map. At each sliding-window location by using 9 anchors multiple region proposals are predicted [5, 13].

When all set of region proposals are generated then it is fed into the CNN to extract a feature vector for each region. Then the feature vectors are used as the input for linear SVMs to predict the class score of each region proposals. At the same time, bounding box coordinates are obtained from the feature vectors by using bounding box regressor.

### 3.8 YOLO

YOLO is used for real time object detection. It is faster than any other object detection algorithm [22, 23]. It does not take whole image as an input for computing the region proposal. Primarily, an image is divided into well defined small grids, then each grid

is bounded by the bounding box. Each bounding box has class score probability if it is above a threshold then it is used to locate the object within the image. Here, a single convolutional network is used to predict the bounding boxes and the class score probabilities of the boxes. YOLO shows lower accuracy in comparison to RCNN but is almost 40 frame per second faster in test accuracy. It faces a problem while detecting small objects within the image.

### ***3.9 Recurrent Neural Networks (RNN)***

The basic network architecture does not support processing of sequential data. This is because, it neither support feature sharing in input data nor work properly for different length of input and output data [27, 30]. Thus, to deal with the sequential dataset, RNN is introduced. RNNs are basically perceptrons having capability to store information about the past events in the hidden layers. It can also update the information of the hidden states in a dynamic way. RNNs only use the information that are earlier used in sequence for prediction. In practice RNNs are not capable to handle long term dependencies.

To overcome from this issue, Long Short Term Memory (LSTM) network is developed. It is a type of RNNs and is used to handle long term dependencies in the network of deep learning architecture. It performs well for sequenced data such as time-series, text generation, NLP, language translation, speech recognition [8], and image captioning etc [14, 26].

### ***3.10 SegNet***

Image segmentation problem is solved in deep learning using SegNet architecture [34]. This architecture is designed by stacking encoder followed by decoder for the task of pixelwise classification within the image. The encoder uses convolution layer, batch normalisation, Relu activation function with non overlapping maxpooling, and sub-sampling. It uses max-pooling indices in the decoders to upsample the low resolution feature maps. This retains the high frequency details in the segmented images and also reduces the total number of trainable parameters in the decoders. The training parameters are tuned by using SGD.

### ***3.11 Generative Adversarial Networks (GANs)***

In GANs two neural networks are used in way of a zero-sum game [7]. One network is working as an adversary to another. It is used to reconstruct or generate the image which is not present in the given dataset. Here, one neural network is working as a



generator and another one as a discriminator. Generator tries to mimic the distribution of real data so that it can mislead the discriminator as shown in Fig. 3. Discriminator does the job of discriminating fake data from the real one [3, 16, 18], both the networks update themselves during the whole process of generation of fake data and their detection. The cost function of both the network are optimized at the same time. So, it is hard to obtain the nash equilibrium.

### 3.12 Capsule Network

A capsule network is a recent deep learning architecture. It is developed with the idea to mimic the human vision system. This network solves the issue that exists with the basic neural network i.e., CNN. In the traditional object detection method, the network performs the detection operation in an invariance way which is not suitable for object detection. It is because the traditional method does not solve the transformation to the image, such as rotation, change in color, or lighting condition. It gives the same object detection result for the transformed image also. It does not support equivariance. Therefore, the capsule network is introduced which support equivariance [25]. It drops the pooling concept to retain all the information about the presence of an object and their orientation by the activity vector. At the primary stage, the initial activity vector goes under affine transformation. After the transformation weighted sum is estimated and finally by using the squash function output vector is obtained.

In the capsule network first 3 layers are encoder layer and the next 3 are decoder [16]. Each decoder has two conv layer and a fully connected layer. The decoder is responsible for reconstructing the original image.

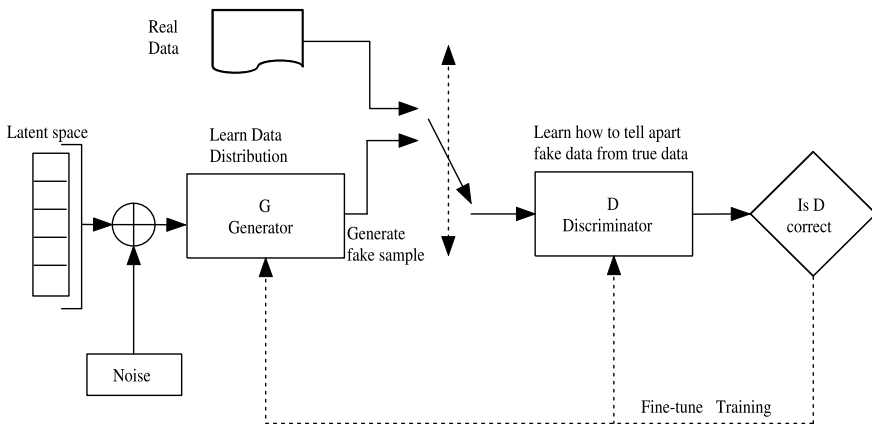


Fig. 3 Working of generative adversarial network [7]

## 4 Experimental Results

Deep learning has applied in various domains and is highly useful in the area of the computer vision. Several authors are using CNN as a base model for image classification [19, 20, 29, 32, 35], image recognition [12], objects detection in an image [13, 22, 36], image segmentation [1, 34], and image generation [7, 37]. The deep neural

**Table 1** Summary of different architectures used in computer vision

Paper	Task	Contribution	Architecture	Dataset
[19]	Image classification	Basic deep neural network model	AlexNet	ILSVRC-2010
[35]	Image classification	Support visualization of intermediate feature layers	ZF net	Caltech-101 and Caltech-256
[29]	Image classification	Increases the depth of neural network with $3 \times 3$ convolution filters	VGG Net	ILSVRC-2012
[20]	Image classification	It deal with the variability of 2D shapes in images	CNN	MNIST
[32]	Image classification	Scale the network and minimize the computational cost	GoogLeNet	ILSVRC 2012
[25]	Classification	It support vector representation for the parameters	Capsule Network	MNIST and CIFAR10
[12]	Image recognition	Skip connection is added for every two layer	Microsoft ResNet	ILSVRC and COCO 2015
[24]	Object detection	Provide object bounds and objectness scores at each position of object	RCNN	PASCAL VOC 2007
[22]	Object detection	Extremely fast for object detection	YOLO	VOC 2012
[13]	Object detection	Add binary mask to Faster-RCNN to generate high-quality segmentation	Mask-RCNN	COCO
[36]	Object detection	Multi-Level Feature Pyramid Network used to construct more effective feature pyramids for detecting objects	M2Det	MS-COCO
[30]	Classification	Support classification of sequential data	RNN	TIMIT database
[34]	Segmentation	Semantic pixel-wise segmentation	SegNet	SUN RGB-D and CamVid
[1]	Segmentation	Atrous spatial pyramid pooling module is used	DeepLabv3	PASCAL VOC 2012
[7]	Image generation	Generate the images that are not available in given dataset	GAN	MNIST, TFD, and CIFAR-10
[37]	Image generation	Learn the natural image manifold directly from data in near real time	iGAN	MIT Places dataset

**Table 2** Different deep learning models for classification of images

Architecture	Test accuracy (%)
AlexNet	84.60
ZF Net	85.40
VGG-16	92.70
GoogLeNet	93.30
ResNet	96.40
CNN	98.88

network architectures used by the above authors are AlexNet, ZF Net, GoogleNet, Capsule network, etc. Table 1 summarises the various architectures used so far in computer vision using deep learning. Table 1 also illustrates the major contributions of authors, dataset they are using for computer visions task. The dataset used by different authors includes ILSVRC-2010, ILSVRC-2012, COCO 2015, TIMIT, MNIST, etc.

The accuracy while the selection of model is crucial for computer vision, in case of security threatening applications [2]. Table 2 details the accuracy achieved by various models used in deep learning. If accuracy of a model is high, its complexity is also higher, so the models are selected as per the requirement or sensitivity of the applications they are going to handle.

## 5 Conclusion

In this paper, we have presented a comprehensive survey of various research challenges in the architectural design of deep learning models and their applications in the field of computer vision. We demonstrate the deep learning models for the class of the problem of image classification, object detection, and image generation. Based upon the various tasks in computer vision the study focuses on the issues and challenges in the architectural design of deep learning models. This comprehensive survey clearly identifies the several research possibilities in deep learning architecture, although several works had been done in deep learning architecture but still a lot more things to be uncovered.

**Acknowledgements** This work is supported by Science and Engineering Research Board (SERB) file number ECR/2017/002419, project entitled as A Robust Medical Image Forensics System for Smart Healthcare, and scheme Early Career Research Award.

## References

1. Chen, L. C., Papandreou, G., Schroff, F., & Adam, H. (2017). *Rethinking atrous convolution for semantic image segmentation*. CoRR abs/1706.05587.
2. Choe, J. W., Nikoozadeh, A., & Oralkan, O., Khuri-Yakub, B.T. (2013). GPU-based real-time volumetric ultrasound image reconstruction for a ring array. *IEEE Transactions on Medical Imaging*, 32(7), 1258–1264.
3. Choi, Y., Choi, M., Kim, M., Ha, J., Kim, S., & Choo, J. (2017). *StarGAN: Unified generative adversarial networks for multi-domain image-to-image translation*. CoRR abs/1711.09020.
4. Forsyth, D. A., & Ponce, J. (2002). *Computer vision: A modern approach*. Pearson Education India.
5. Girshick, R. (2015). Fast R-CNN. In *Proceedings of IEEE ICCV* (pp. 1440–1448).
6. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
7. Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). Generative adversarial nets. In *Proceedings of NIPS* (pp. 2672–2680).
8. Graves, A., Mohamed, A., & Hinton, G. (2013). Speech recognition with deep recurrent neural networks. In *Proceedings of IEEE ICASSP* (pp. 6645–6649).
9. Guo, T., Dong, J., Li, H., & Gao, Y. (2017). Simple convolutional neural network on image classification. In *Proceeding of IEEE ICBDA* (pp. 721–724).
10. Hall, M. A., & Smith, L. A. (1999). Feature selection for machine learning: Comparing a correlation-based filter approach to the wrapper. In *Proceedings of IFAIRSC* (pp. 235–239).
11. Hatcher, W. G., & Yu, W. (2018). A survey of deep learning: Platforms, applications and emerging research trends. *IEEE Access*, 6, 24411–24432.
12. He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep residual learning for image recognition*. CoRR abs/1512.03385.
13. He, K., Gkioxari, G., Dollr, P., & Girshick, R. (2017). Mask R-CNN. In *Proceedings of IEEE ICCV* (pp. 2980–2988).
14. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780.
15. Hu, B., Lu, Z., Li, H., & Chen, Q. (2014). Convolutional neural network architectures for matching natural language sentences. *Advances in Neural Information Processing Systems*, 27, 2042–2050.
16. Jaiswal, A., AbdAlmageed, W., & Natarajan, P. (2018). CapsuleGAN: Generative adversarial capsule network. In *ECCV Workshops*.
17. Kaïming, H., Zhang, X., Ren, S., Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imageNet classification. In *Proceedings of IEEE ICCV* (pp. 1026–1034).
18. Karras, T., Laine, S., & Aila, T. (2018). *A style-based generator architecture for generative adversarial networks*. CoRR abs/1812.04948.
19. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 1097–1105.
20. Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324.
21. O’Shea, K., & Nash, R. (2015). *An introduction to convolutional neural networks*. ArXiv e-prints.
22. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016a). You only look once: Unified, real-time object detection. In *Proceeding of IEEE CVPR* (pp. 779–788).
23. Redmon, J., Divvala, S. K., Girshick, R. B., & Farhadi, A. (2016b). You only look once: Unified, real-time object detection. *Proceeding of IEEE CVPR* (pp. 779–788).
24. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, 28, 91–99.
25. Sabour, S., Frosst, N., & Hinton, G. E. (2017). *Dynamic routing between capsules*. CoRR abs/1710.09829, 1710.09829.

26. Sak, H., Senior, A. W., & Beaufays, F. (2014). *Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition*. CoRR abs/1402.1128.
27. Schuster, M., & Paliwal, K. (1997). Bidirectional recurrent neural networks. *Transaction in Signal Processing*, 45(11), 2673–2681.
28. Shin, H., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., et al. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging*, 35(5), 1285–1298.
29. Simonyan, K., & Zisserman, A. (2014). *Very deep convolutional networks for large-scale image recognition*. CoRR abs/1409.1556.
30. Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in Neural Information Processing Systems*, 27, 3104–3112.
31. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., et al. (2015). Going deeper with convolutions. In *Proceeding of IEEE CVPR* (pp. 1–9).
32. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceeding of IEEE CVPR*.
33. Turner, C. R., Wolf, A. L., Fuggetta, A., & Lavazza, L. (1998). Feature engineering. In *Proceedings of IWSSD* (p. 162).
34. Vijay, B., Kendall, A., & Cipolla, R. (2017). SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
35. Zeiler, M. D., & Fergus, R. (2013). *Visualizing and understanding convolutional networks*. CoRR.
36. Zhao, Q., Sheng, T., Wang, Y., Tang, Z., Chen, Y., Cai, L., et al. (2018). *M2Det: A single-shot object detector based on multi-level feature pyramid network*. CoRR abs/1811.04533.
37. Zhu, J. Y., Krähenbühl, P., Shechtman, E., & Efros, A. A. (2016). *Generative visual manipulation on the natural image manifold*. CoRR abs/1609.03552.

# Robust Reversible Watermarking for Grayscale Medical Images



Tanima Dutta, Randheer Bagi and Hari Prabhat Gupta

**Abstract** Robust reversible watermarking algorithms applied to protect medical imaging from misdiagnosis for any slight distortion. It suited for tasks like copyright protection while preserving the visual quality of watermarked images. It also shows robustness against transmission errors. We propose a novel robust reversible watermarking technique for grayscale medical images. We investigate the high embedding capacity watermarking technique also keeping low distortion in the watermarked images. We demonstrate the accuracy of the reversibility even in error-prone transmission channels. Our simulations show that the proposed technique effectively retains the perceptual quality.

**Keywords** Reversible watermarking · Grayscale images · Medical imaging · Robust watermarking · Hardware efficient

## 1 Introduction

Reversible watermarking is popular in medical imaging where even slight distortion in images is not affordable because such distortion may lead to misdiagnosis. Reversible watermarking embeds a watermark into digital images in a reversible fashion. At the receiver end, the original image is recovered in a lossless way. This property is notably beneficial to preserve the image quality although the watermark is embedded to protect its authenticity and integrity. Medical imaging, multimedia archive, digital security in defense, and remote sensing are the areas in which

---

T. Dutta (✉) · R. Bagi · H. P. Gupta  
Department of Computer Science and Engineering,  
IIT (BHU) Varanasi, Varanasi, India  
e-mail: [tanima.cse@iitbhu.ac.in](mailto:tanima.cse@iitbhu.ac.in)

R. Bagi  
e-mail: [randheerbagi.rs.cse17@iitbhu.ac.in](mailto:randheerbagi.rs.cse17@iitbhu.ac.in)

H. P. Gupta  
e-mail: [hariprabhat.cse@iitbhu.ac.in](mailto:hariprabhat.cse@iitbhu.ac.in)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_57](https://doi.org/10.1007/978-981-15-0694-9_57)

reversible watermarking widely used. In practice, the conventional reversible watermarking methods are not actively applicable where watermarked data are transmitted over networks. The reasons can be summarized as follows:

- The ubiquity of transmission errors over the network makes extraction of watermarks fragile in nature;
- Sometimes transmission error may cause the disappearance of watermarked images;
- These are not at all suitable for low-cost hardware;
- Obtaining agreeable reversibility for huge image dataset is difficult;

Therefore, it is highly required to address these issues by developing a framework. Robust reversible watermarking ensures the recovery of host images and watermarks over the lossless channel. It also counters distortions in the noised channels. The robust reversible watermarking is thus a challenging task. Less amount of work has been done in literature to address this problem [5–14]. A hardware efficient technique can be implemented efficiently in low-cost hardware if it poses lower runtime complexity [9]. The designing of a robust reversible watermarking framework in inexpensive device demands these four objectives: (1) invisibility, (2) reversibility, (3) robustness, and (4) runtime complexity.

**Contribution:** In this paper, we focus on grayscale images for high embedding capacity of the watermark in a host image with the high visual quality of the watermarked image. We also direct a distortion-free robust reversible watermarking technique. The main contributions are compiled as follows:

- In each grayscale medical image, first the nonoverlapping  $4 \times 4$  pixel blocks, where correlation amount the pixels, are very high which are selected. All consecutive pixels are paired to form nonoverlapping pairs in a block. A watermark sequence is invisibly embedded in these pairs.
- Next, a reversible watermarking algorithm is proposed with blind extraction method. The algorithm is robust against lossless and lossy compression methods and common image processing attacks.
- Then, an extraction algorithm is presented which can recover the original pixels at the receiver even in presence of transmission errors.
- Last, we have evaluated perceptual quality and robustness of the propose watermarking technique with simulations.

The proposed technique offers a lossless and robust watermarking technique. It counters the embedding and channel distortion to assure the visual quality of the watermarked image for the human visual system (HVS).

The remainder of this paper is regulated as follows: In Sect. 2, we quickly discuss the literature review. The proposed novel robust reversible watermarking technique is discussed in Sect. 3 with its analysis and validation by simulation. Section 4 shows the results of simulations and performance evaluation of the reversible algorithm for different grayscale medical images in both lossy and lossless environment. At last, we conclude our paper in Sect. 5.

## 2 Literature Survey

Reversible watermarking popularly applied in the application field where we need to recover the original data after executing the task. Many reversible watermarking methods are available in the literature for images [15–17, 23], where the authors assumed that the transmission channel is lossless. In practice, the transmission channel is not entirely lossless. So a robust reversible watermarking technique is needed to deal with losses in the transmission channel. In literature review, various reversible watermarking techniques are introduced to implement watermarking in video compressed domain [6]. This compression technique ended more compressed information about the video for watermarking without hampering its perceptual quality. In various applications [5, 7, 8, 10, 12] the H.264/AVC compression technique are used for video watermarking.

In the histogram rotation-based methods [4, 25], the centroid vectors of two random zones rotates in the nonoverlapping blocks. These methods were sensitive to conventional image processing attacks. Histogram distribution constrained-based methods are used to resolve these issues, in spatial, and wavelet domains in literature [19, 26]. In this method, an image divided into different types of blocks. The watermarking is performed in each block type by using histogram distribution.

Simplicity and stability are two properties supported by statistical quantity histogram which gains the immense attraction of researchers. Some of the examples are difference histogram [21], the arithmetic average of difference histogram [1], and prediction error histogram [13]. The generalized statistical quantity histogram method proposed in [14]. In this, embedding and extraction of a watermark achieved in the wavelet coefficients by using histogram shifting. The method as mentioned earlier for watermarking is extended in [2] by using histogram shifting and clustering. Moreover, to satisfactory balance robustness and invisibility, the authors have used the enhanced pixel-wise masking of literature [3].

**Motivation:** The key observations of the literature motivate this paper. In literature [4, 25] the author reported a robust watermarking technique which is against to lossy compressions, but it is sensitive to common image processing attacks. It leads to poor visual quality of watermarked images also affects the restoration of host images. The methods proposed in literature [19, 26] experience weak reversibility and robustness [2]. In literature [1, 2, 13, 14, 22], the robustness is achieve at the cost of high computational complexity. Such algorithms are not practically applicable for medical data transmission. In short, the above analysis shows that existing robust reversible watermarking methods are not easily applicable in practice.

## 3 Proposed Watermarking Technique

A host image is divided into  $4 \times 4$  pixel blocks. The pixels are scanned in a  $4 \times 4$  pixel blocks in zigzag sequence order. The proposed watermarking technique embeds a watermark sequence invisibly in these pairs.



In this section, first we propose a watermark embedding algorithm, next the extraction of watermark bits with the recovery of original pixels in both lossless and the lossy environment is presented, and finally, the complexity analysis of the embedding and extraction algorithms is described. The locations of embedding watermark bits are maintained in location map (a binary file).

### 3.1 Pair Selection

Assume the size of an image, number of  $4 \times 4$  pixel blocks, and number of pairs in a pixel block are  $m \times n$ ,  $N$ , and  $M$ , respectively. Two consecutive pixels in the sequence are paired together.  $A_{ij}$  and  $B_{ij}$  denote 1st and 2nd pixel in  $i$ th pair ( $A_{ij}, B_{ij}$ ) of  $j$ th block, respectively, where  $0 \leq i \leq M$  and  $0 < j \leq N$ . The difference between the pixels in  $i$ th pair of  $j$ th block, denoted by  $d_{ij}$  is expressed by

$$A_{ij} - B_{ij} = d_{ij} \text{ and } |d_{ij}| = T_{ij}, \quad (1)$$

where  $T_{ij}$  denotes the absolute difference between the pixels in  $i$ th pair of  $j$ th block of the image. A pair ( $A_{ij}, B_{ij}$ ) is chosen for watermark embedding as per the following function:

$$S(i, j) = \begin{cases} 1 & \text{if } (d_{ij} > 0 \text{ and } B_{ij} > 2T_{ij}) \\ 1 & \text{if } (d_{ij} < 0 \text{ and } A_{ij} > 2T_{ij}) \\ 1 & \text{if } (d_{ij} = 0) \\ 0 & \text{otherwise} \end{cases}, \quad (2)$$

where the value of function  $S(i, j)$  as 1 indicates that  $i$ th pair of  $j$ th block is selected for watermark embedding. The embedding distortion induced the perceptual quality degradation is also minimized using the Eq. 2, where a limit is applied in selecting pairs so that the value of any pixel will never be increased to its double for embedding a watermark bit. Each pair in location map requires two bits. The  $i$ th pair of  $j$ th block in the location map is denoted by  $map_{ij}$ . If  $map_{ij} = 00$ , then  $i$ th pair of  $j$ th block is not selected for watermark embedding. If  $map_{ij} \neq 00$ , then watermark can be embedded in that pair. The number bits in the location map are  $m \times n$ . Therefore, the maximum embedding capacity can express as

$$N \leq \frac{m \times n}{2}. \quad (3)$$

### 3.2 Watermark Embedding

A random watermark sequence is embedded in each selected pair of pixel blocks of an image in invisible fashion. The watermark which is going to embed is binary (0, 1) in nature. The  $j$ th pixel block in the location map, denoted by  $map_j$  is initialized

with zero values. If a watermark bit, denoted by  $W_{ij}$ , is 0 then  $\bar{A}_{ij} = A_{ij} + 2D_{ij}$  and  $\bar{B}_{ij} = B_{ij} - 2D_{ij}$ , otherwise  $\bar{B}_{ij} = B_{ij} + 2D_{ij}$  and  $\bar{A}_{ij} = A_{ij} - 2D_{ij}$ . The value of  $D_{ij}$  and  $map_{ij}$  however depend on the value of  $d_{ij}$  and  $W_{ij}$ , which can be given by

- if  $d_{ij} < 0$  and  $W_{ij}$  is 0, then  $D_{ij} = |d_{ij}|$  and  $map_{ij}=11$ ;
- if  $d_{ij} < 0$  and  $W_{ij}$  is 1, then  $D_{ij} = 0$  and  $map_{ij} = 01$ ;
- if  $d_{ij} > 0$  and  $W_{ij}$  is 0, then  $D_{ij} = 0$  and  $map_{ij} = 01$ ;
- if  $d_{ij} > 0$  and  $W_{ij}$  is 1, then  $D_{ij} = |d_{ij}|$  and  $map_{ij}=11$ ;
- if  $d_{ij} == 0$ , then  $D_{ij} = 1$  and  $map_{ij} = 10$ .

### 3.3 Reversibly Extract and Authenticate

Extraction of watermark sequence is performed by the decoder. It is a reverse process of embedding. The watermarked image is divided into nonoverlapping  $4 \times 4$  pixel blocks. Pixels are scanned in zigzag sequence. The 1st and 2nd pixels in  $i$ th pair of  $j$ th block of the image are denoted by  $\bar{A}_{ij}$  and  $\bar{B}_{ij}$ , respectively. If  $\bar{A}_{ij} > \bar{B}_{ij}$ , then watermark bit is 0, otherwise 1. The watermarked pixels ( $\bar{A}_{ij}$  and  $\bar{B}_{ij}$ ) are recovered to its original value ( $A_{ij}$  and  $B_{ij}$ ) based on the value of location map. If the value of location map for a pair is 00, then the pair is not watermarked so reversibility is not required.

- If  $map_{ij} = 11$ , then  $A_{ij} = \bar{A}_{ij} - (2/3)\bar{D}_{ij}$  and  $B_{ij} = \bar{B}_{ij} + (2/3)\bar{D}_{ij}$  for  $\bar{A}_{ij} > \bar{B}_{ij}$  and  $B_{ij} = \bar{B}_{ij} - (2/3)\bar{D}_{ij}$  and  $A_{ij} = \bar{A}_{ij} + (2/3)\bar{D}_{ij}$  for  $\bar{A}_{ij} < \bar{B}_{ij}$ , where  $\bar{D}_{ij} = |\bar{A}_{ij} - \bar{B}_{ij}|$ .
- Similarly, if  $map_{ij} = 10$ , then  $A_{ij} = \bar{A}_{ij} - 2\bar{D}_{ij}$  and  $B_{ij} = \bar{B}_{ij} + 2\bar{D}_{ij}$  for  $\bar{A}_{ij} > \bar{B}_{ij}$  and  $B_{ij} = \bar{B}_{ij} - 2\bar{D}_{ij}$  and  $A_{ij} = \bar{A}_{ij} + 2\bar{D}_{ij}$  for  $\bar{A}_{ij} < \bar{B}_{ij}$ , where  $\bar{D}_{ij} = 1$ .
- On the other hand, if  $map_{ij} = 01$  the value of watermarked pixels and original pixels are same.

### 3.4 Authenticity and Reversibility in Error-Prone Environment

Any transformations (like blurring, geometric transformation, or lossy compressions) can easily destroy least significant bit (LSB) of pixels as mentioned in literature [20]. The proposed watermarking technique can recover pixels in such error-prone environment. However, the recovery of pixels depends on the value of the location map. The extraction of watermark bit and restoration of pixels in error-prone environment when  $map_{ij} = 10$ , where  $\bar{T}_{ij}$  is the absolute difference between the watermarked pixels in a pair.

Similarly, a watermark bit is extracted and the pixels are recovered when  $map_{ij} = 11$ . The original pixels cannot be recovered from watermarked and unwatermarked pixels in error-prone environment when  $map_{ij} = 01$  and  $map_{ij} = 00$ , respectively.

However, when  $map_{ij} = 01$  then the watermark bit can be extracted correctly in error-prone environment unless the sign of the difference between watermarked pixels is flipped or the difference becomes zero.

## 4 Performance Evaluation

The proposed watermarking technique is simulated using datasets [18] and performance is evaluated. It is a free open-access online database of medical images. It has grayscale radiological images in the database over 12,000 patients. Dataset has more than 59,000 indexed and curated images. The images in the dataset are structured in different categories like disease location (organ system); pathology; profiles of patients; and image captions. The dataset is searchable by patient symptoms and signs, diagnosis, organ system, image modality and image description, keywords, contributing authors, and many other search options. Figure 1 shows nine radiology images which are used for experimentation. The two metrics, Peak Signal-to-Noise Ratio and Bit Error Rate, are used for performance evaluation. It also ensures the visual quality and robustness of the watermarked images.

### 4.1 Peak Signal-to-Noise Ratio (PSNR)

After extraction of the watermark sequence the quality of the reconstructed image is evaluated by peak signal-to-noise ratio [24] such that

$$\text{PSNR}(dB) = 10 \log \frac{255^2}{\langle (a_i - \bar{a}_i)^2 \rangle}, \quad (4)$$

where  $a_i$  and  $\bar{a}_i$  are pixel values in the host and recovered images, respectively, and the averaging operator is denoted by  $\langle \cdot \rangle$ .

Figure 2 depicts the quality of recovered images using the proposed method against salt and pepper noise and gaussian noise for the images shown in Fig. 1.

### 4.2 Bit Error Rate (BER)

The frequency of bit errors at the time of detecting a multi-bit watermark message is known as the Bit Error Rate [11], such that

$$\text{BER} = \frac{\text{number of error bits in watermarked video stream}}{\text{total bits sent}} \quad (5)$$



Fig. 1 Illustration of radiological images used for experimentation [18]

and robustness of the proposed reversible watermarking technique is expressed as

$$\text{Robustness} = (1 - \text{BER}) \times 100. \tag{6}$$

Figure 3 depicts that robustness of the proposed technique against salt and pepper noise and gaussian noise for the images shown in Fig. 1.

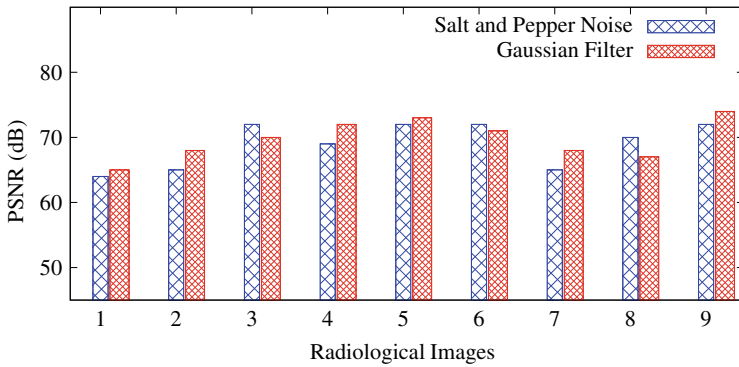


Fig. 2 Peak Signal-to-Noise Ratio (PSNR) against image processing attacks

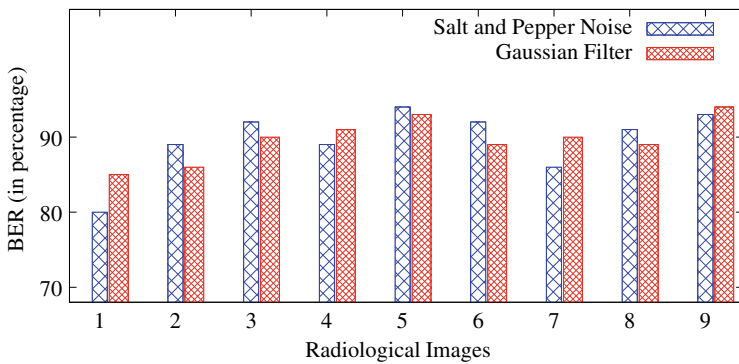


Fig. 3 Bit Error Rate (BER) for image processing attacks

## 5 Conclusion

We proposed a spatial domain-based novel reversible watermarking algorithm with blind extraction for grayscale medical images. It has focused on highly correlated pixels and exploited their spatial redundancy to achieve improved performance in terms of visual quality and robustness. Low complexity algorithms are proposed so that the watermarking technique can be implemented in low-cost hardware. Performance evaluation performed by the metrics like reversibility, robustness, invisibility, capacity, and runtime complexity. The proposed technique is readily applicable in practice.

We consider that this paper stimulates researchers for further research in reversible watermarking to improve robustness and particularly to reduce embedding distortion of watermarked medical images.

**Acknowledgements** This work is supported by Science and Engineering Research Board (SERB) file number ECR/2017/002419, project entitled as A Robust Medical Image Forensics System for Smart Healthcare, and scheme Early Career Research Award.

## References

1. An, L., Gao, X., Deng, C., & Ji, F. (2009). Reversible watermarking based on statistical quantity histogram. In *Proceedings of PCM: Advances in Multimedia Information Processing*, pp. 1300–1305.
2. An, L., Gao, X., Li, X., Tao, D., Deng, C., & Li, J. (2012). Robust reversible watermarking via clustering and enhanced pixel-wise masking. *IEEE Transactions on Image Processing*, 21(8), 3598–3611.
3. Barni, M., Bartolini, F., & Piva, A. (2001). Improved wavelet-based watermarking through pixel-wise masking. *IEEE Transactions on Image Processing*, 10(5), 783–791.
4. De Vleeschouwer, C., Delaigle, J., & Macq, B. (2003). Circular interpretation of bijective transformations in lossless watermarking for media asset management. *IEEE Transactions on Multimedia*, 5(1), 97–105.
5. Dutta, T. (2013). Motion compensated compressed domain watermarking. In *Proceedings of ACM International Conference on Multimedia*, pp. 1039–1042.
6. Dutta, T. (2015). Medical data compression and transmission in wireless Ad Hoc networks. *Sensors Journal, IEEE*, 15, 778–786.
7. Dutta, T., & Gupta, H. P. (2016). A robust watermarking framework for High Efficiency Video Coding (HEVC) Encoded video with blind extraction process. *Journal of Visual Communication and Image Representation*, 38, 29–44.
8. Dutta, T., & Gupta, H. P. (2017). An efficient framework for compressed domain watermarking in P frames of high-efficiency video coding (HEVC)-encoded video. *ACM Trans Multimedia Computer Communications Applications*, 13(1), 12:1–12:24.
9. Dutta, T., & Gupta, H. P. (2017). Leveraging smart devices for automatic mood-transferring in real-time oil painting. *IEEE Transactions on Industrial Electronics*, 64(2), 1581–1588.
10. Dutta, T., Sur, A., & Nandi, S. (2013a). A robust compressed domain video watermarking in P-frames with controlled bit rate increase. In *Proceedings of National Conference on Communications (NCC)*, pp. 1–5.
11. Dutta, T., Sur, A., & Nandi, S. (2013b). A robust compressed domain video watermarking in P-frames with controlled bit rate increase. In *Proceeding of NCC*, pp. 1–5.
12. Dutta, T., Sur, A., & Nandi, S. (2013c). MCRD: Motion coherent region detection in H.264 compressed video. In *Proceedings of IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6.
13. Fallahpour, M. (2008). Reversible image data hiding based on gradient adjusted prediction. *IEICE Electronics Express*, 5(20), 870–876.
14. Gao, X., An, L., Yuan, Y., Tao, D., & Li, X. (2011). Lossless data embedding using generalized statistical quantity histogram. *Circuits and Systems for Video Technology, IEEE Transactions on*, 21(8), 1061–1070.
15. Kim, K., Lee, M., Lee, H., & Lee, H. (2009). Reversible data hiding exploiting spatial correlation between sub-sampled images. *Pattern Recognition*, 42(11), 3083–3096.
16. Lee, S., Yoo, C., & Kalker, T. (2007). Reversible Image Watermarking Based on Integer-to-Integer Wavelet Transform. *IEEE Transactions on Information Forensics and Security*, 2(3), 321–330.
17. Lin, C., Tai, W., & Chang, C. (2008). Multilevel reversible data hiding based on histogram modification of difference images. *Pattern Recognition*, 41(12), 3582–3591.
18. Medicine TNLo (2019) MedPix. <https://medpix.nlm.nih.gov/home>.

19. Ni, Z., Shi, Y., Ansari, N., Su, W., Sun, Q., & Lin, X. (2008). Robust lossless image data hiding designed for semi-fragile image authentication. *IEEE Transactions on Circuits and Systems for Video Technology*, 18(4), 497–509.
20. Said, A., & Pearlman, W. (1996). An image multiresolution representation for lossless and lossy compression. *IEEE Transactions on Image Processing*, 5(9), 1303–1310.
21. Tai, W. L., Yeh, C. M., & Chang, C. C. (2009). Reversible data hiding based on histogram modification of pixel differences. *IEEE Transactions on Circuits and Systems for Video Technology*, 19(6), 906–910.
22. Thabit, R., & Khoo, B. E. (2014). Robust reversible watermarking scheme using slantlet transform matrix. *Journal of Systems and Software*, 88, 74–86.
23. Tian, J. (2003). Reversible data embedding using a difference expansion. *IEEE Transactions on Circuits and Systems for Video Technology*, 13(8), 890–896.
24. Turcza, P., & Duplaga, M. (2013). Hardware-efficient low-power image processing system for wireless capsule endoscopy. *IEEE Journal of Biomedical and Health Informatics*, 17(6), 1046–1056.
25. Vleeschouwer, C., Delaigle, J., & Macq, B. (2001). Circular interpretation of histogram for reversible watermarking. In *Proceeding of IEEE Workshop on MSP*, pp. 345–350.
26. Zou, D., Shi, Y., Ni, Z., & Su, W. (2006). A semi-Fragile lossless digital watermarking scheme based on integer wavelet transform. *IEEE Transactions on Circuits and Systems for Video Technology*, 16(10), 1294–1300.

# Improved Detection of Kidney Stone in Ultrasound Images Using Segmentation Techniques



Rati Goel and Anmol Jain

**Abstract** The medical images are often corrupted by various noises and blurriness. In particular, the noises presented in ultrasound images may lead to an inaccurate diagnosis of smaller kidney stones and affect its treatment. This paper proposes an improved technique for detection of kidney stone from the ultrasound images of kidney. The ultrasound kidney images are preprocessed to remove labels and change from RGB to Gray images. Further, image contrast is enhanced by adjusting the image intensity. To remove the noises, median filtering is used. The filtered image is taken as input for morphological segmentation process; initially applied dilation and then seed region growing algorithm is used to segment the renal calculi from ultrasound image of kidney. The region parameters are extracted from the segmented region. Finally, area of each renal calculi is calculated. The various performance evaluation GLCM features such as entropy, contrast, angular second moment and correlation are used to judge the quality of output images. The confusion matrix is also prepared to analyze the sensitivity, specificity, and accuracy of the final system. The overall accuracy of classification system is around 90%. The proposed technique may help medical professionals in easy detection of kidney stones and benefit the patients.

**Keywords** Kidney · Region of interest · Speckle noise reduction · Confusion matrix · Segmentation · Filters · Morphological operation · Ultrasound

## 1 Introduction

The medical images such as Magnetic resonance imaging (MRI), Computed tomography (CT), and ultrasound images are widely used for diagnosis, effective planning, and clinical research. Medical image has become almost compulsory to help the radiologists by using digital computer images for accurate clinical diagnosis [1].

---

R. Goel (✉) · A. Jain  
ABES Engg. College, Ghaziabad, India  
e-mail: [rati.sept@gmail.com](mailto:rati.sept@gmail.com)

A. Jain  
e-mail: [anmol.jain@abes.ac.in](mailto:anmol.jain@abes.ac.in)

© Springer Nature Singapore Pte Ltd. 2020  
M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_58](https://doi.org/10.1007/978-981-15-0694-9_58)



Ultrasound imaging techniques play a crucial role in emergency diagnostic method. It is widely used due to its non-invasive inexpensive availability and non-radiation exposure [2]. Various types of kidney diseases are listed under chronic Kidney diseases which may cause severe health problem. This work emphasizes the most prevalent diseases occurring in the kidney region for detection. It mainly emphasized on kidney stone, kidney cyst, and renal cell cancer [3].

Medical images are reconstruction for a deeper understanding of clinical abnormalities including brain tumor, breast cancer, and kidney stone disease, etc. Kidney stone disease is one of the major life intimidating disorders continuing world wide. The kidney diseases can be grouped into two main stages namely acute kidney injury (AKI) and chronic kidney diseases [4]. The prevalence of chronic kidney diseases will gradually increase if they are not properly treated. It may initiate serious health hazards namely diabetes, blood pressure, pulmonary hypertension, and other cardiovascular diseases. Kidney has two basic functions—disposing toxic substances from the blood and preserves the useful components in proper balance. According to the statistics of National Centre for Biotechnology Information (NCBI) there is 30% increase in the prevalence of chronic kidney diseases in United States, In India 40–60% of diabetic and hypertension cases are due to CKD [5]. It is more necessary to diagnose the kidney diseases at early stages which can prevent us from several serious diseases. Ultrasound modality is one of the best imaging diagnostic techniques when compared to other imaging modalities such as Computed tomography (CT), X-ray and Magnetic resonance imaging (MRI) because of it is available at less expense with no harmful radiation exposure and its smart portability. Various types of kidney diseases are listed under chronic Kidney diseases which may cause severe health problem. This work emphasizes the most prevalent diseases occurring in the kidney region for detection. It mainly emphasized on kidney stone, kidney cyst, and renal cell cancer [6]. A number of methods have been used for diagnosing the kidney stone such as blood test, urine test, scanning. Scanning also differs in CT scan, Ultrasound scan, and Doppler scan. Nowadays a field of automation came into existence which also being used in medical field. There are many types of tools(computer-assisted) Such as ultrasound, CT scan (computed tomography), and X-rays that deliver the most exact diagnostic tools for kidney stone screening and diagnosis. In image processing method various methods may be performed. However, firstly, an image is converted into a digital form to obtain an enhanced image or to get some relevant information from it.

The Kidney stone detection is a challenging task as ultrasound images have a low-resolution that is poor quality and contain speckle noise. Low quality image makes highly difficult human as well as machine examination. The medical field cannot afford low accuracy as the human life is involved [2]. Thus, the classification techniques need to be used to analyze to improve detection of kidney stone. In this paper, the unwanted labels, texts, and noises from ultrasound images have been removed by median filtering, intensity adjustment, morphological operation followed by the threshold, and region-based segmentation techniques.

This chapter is organized as follows: Sect. 2 presents the literature review, the proposed methodology is discussed in Sect. 3, for analysis purposes the confusion matrix is elaborated in Sect. 4, Sect. 5 depicts the results and conclusion in Sect. 6.

## 2 Literature Review

A novel classifier for segmentation and identification of kidney stone from the ultrasound image has been proposed by Selvarani and Rajendiran [7]. The Discrete Wavelet Transform has been used to eliminate noises and enhanced the image. Further, Chanvase algorithm has been implemented for segmentation. The histogram oriented Gradient and GLCM have been employed for feature extraction to improve accuracy [7]. An ensemble model formulated by texture features using GLCM and GLDM has been created by Bhart et al. [8]. Anjana and Kaur [9] shared the techniques of image segmentation to improve image features such as intensity values of pixels and textures, etc. Filtering techniques (median filter and Gaussian filter) have been presented to enhance the image quality. After that they used morphological operations and then to find the region of interest they used entropy-based segmentation. Finally for the analysis of kidney stone images, they use KNN and SVM classification techniques [10]. The automatic detection of kidney diseases using Viola Jones method incorporated with different features is used in smart phone by [11]. Saroha et al. [6] implemented the codes of geometry for defining the borderline and partitioning the kidney area by using segmentation techniques and enhancing the detection of kidney stone. The ideal, median, and Butterworth filters have been used. The performance of these filters has been analyzed on the basis of MSE, PSNR, and SNR. A clear vision about the identification of kidney stones based on the binary conversion using threshold range and the morphological filtration of the ultrasound image has been presented by Nithyavathy [12].

Hu et al. [13] proposed the speckle reduction on ultrasound images to remove high multiplicative noises created by back scattered waves. The authors have implemented the program to use image processing techniques and codes of geometry to define the borderline and partition the kidney area by using segmentation techniques. This enhances the detection of kidney stone [1]. The segmentation technique is used to detect descriptive multiple stones, processed the wavelets for kidney stone identification, and classification by using ANN et al. [14]. Rahman and Uddin [15] have proposed a technique to the reduction of speckle noise and partition by using segmentation in US image. Tijjani and Sani [16] provide an overview of the ANN-based approaches to predict kidney problem. They designed an algorithm to remove the speckle noise from ultrasound medical images. They also used Mathematical Morphological operations in their algorithm based on Morphological Image Cleaning algorithm [17]. The literature review reveals that a number of techniques such as Discrete Wavelet Transform, filtering, Segmentation, and Morphological techniques have been used. However, these techniques have been implemented on a smaller set of medical images with low degree of accuracy. In the presented work, the accuracy of kidney stone detection has been improved on a larger set of biomedical images.

### 3 Proposed Methodology

This paper has proposed a methodology for identifying the presence of renal calculi (kidney stone) in the renal ultrasound images. The presented method also helps in measuring the size and area of the kidney stones for segmentation process. The proposed technique starts with image acquisition, by this, the ultrasound kidney images are preprocessed to remove labels and change from RGB to Gray images. Further, image contrast is enhanced by adjusting intensity of image. To remove the noises median filtering is used. The filtered image is taken as input for segmentation process; seed region growing algorithm is used to segment the renal calculi from ultrasound image of kidney. The region parameters are extracted from the segmented region [18]. Finally area of each renal calculi is calculated. The various performance evaluation GLCM features such as entropy, contrast, angular second moment, and correlation are used to judge the quality of output images. The confusion matrix is also prepared to analyze the sensitivity, specificity, and accuracy of the final system [19].

The flow chart of the proposed methodology for the detection of kidney stone is depicted in Fig. 1. The major steps of methodology consist of (i) Image Acquisition (ii) Image preprocessing (iii) Image enhancement and filtration (iv) Morphological and segmentation, Region of interest, and Feature Extraction. Using these steps the proposed system gives a robust detection procedure.

#### 3.1 Image Acquisition

The proposed method starts with image acquisition used to take an image from the external source of system. In this method the image is acquired using image acquisition toolbox command in MATLAB. The main role of these toolbox commands is to make the image readable to the machine and can apply other operation on the same provided format. For this work, at first the acquired dicom images are converted into jpeg format for processing.

#### 3.2 Image Preprocessing

In Image preprocess, achieve the original image from the degraded image. Images are preprocessed to remove labels or text inside the given images and change from RGB to Gray images.

Preprocessing always yields better results for further processing. After preprocessing, the image became clear, sharp with removal of labels and text. It can be easily seen in Figs. 2 and 3.

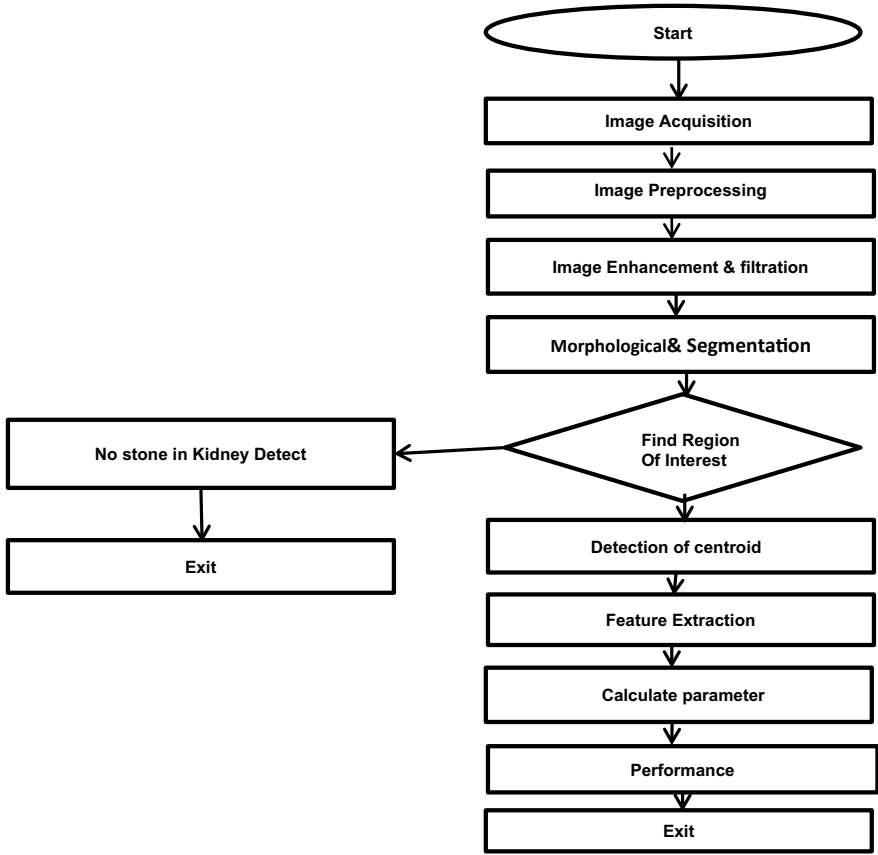
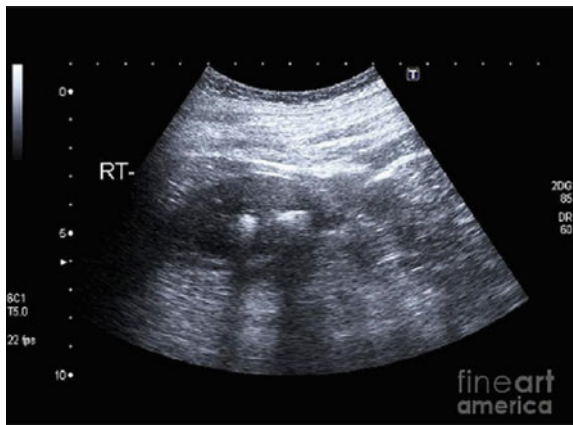


Fig. 1 Flow chart of proposed technique for kidney stone detection

Fig. 2 Input image



**Fig. 3** Preprocessed image



### 3.3 Image Enhancement and Filtration

The main aim of image enhancement is processing on an image in order to make it more appropriate for certain applications. Image enhancement mainly sharpens image features like boundaries, edges or contrast and reduces the ringing artifacts. The enhancement improves the picture quality so that the information contained in them could be extracted in a meaningful sense. Contrast enhancement technique help to increase or decrease the contrast of an image accordingly as contained in Fig. 4. Three methods are particularly used for contrast enhancement: intensity adjustment, histogram equalization, and adaptive histogram equalization. Contrast enhancement technique of the gray scale image is different from the color images [20]. There are many areas such as vision, autonomous navigation, dynamic scene analysis, and biomedical image analysis requiring image modification to show the information in a better way and reveal the important content. Image enhancement techniques are widely used in various areas of engineering and science. Exterior noises and environmental conflicts affect the quality of images by ambient pressure and temperature fluctuations. Thus, image enhancement is necessary. You can apply image processing in every field where images are to be implicit and analyzed. Such as image analysis in medical field, satellite image analysis, etc. By using image enhancement techniques we can modify the image components to increase clarity, sharpness, and details through the visual analysis and interpretation. It appear also to transform the graphical impact in a way that heightens the information parts of the image [6].

**Fig. 4** Submission for image enhancement



Many image enhancement techniques are used to enhance the quality of the digital image without affecting any destruction to it. There are several of the image enhancement procedures which contain enhancement of contrast, saturation transformations, intensity, edge enhancement, hue, and gray level slicing. Usually the ultrasound images have speckle noise due to the high frequency backscattered waves on transducer [21]. The Median filter produces better noise reduction results and preserves edge without much loss of information when compared to other filtering techniques [15]. The median filter is used to obtain an approximation of the original scene.

### 3.4 Morphological and Image Segmentation

Erosion and dilation are morphological operations. To find the boundaries of the image or what we need to detect by using erosion and dilation [22]. The dilation equation is as follows

$$A \oplus B = \{Z | (\hat{B})_Z I A \neq \emptyset\} \tag{1}$$

where,  $\emptyset$  and  $\hat{B}$  are the empty set and reflection of the structuring element B.

The erosion equation is defined as:

$$A \ominus B = \{Z | (B)_Z I A^C \neq \emptyset\} \tag{2}$$

where  $A^C$  is the complement of A.

To detect the stone in the kidney, we will apply the segmentation technique [23]. By using segmentation, we partition an image into distinct areas, that distinct part collectively covers the entire image or set of contours extracted from image [24].

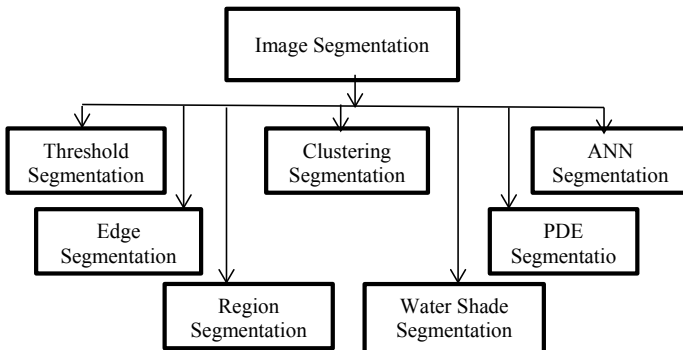


Fig. 5 Image segmentation techniques

As shown in Fig. 5. Using Image Segmentation images can be divided into several segments. Each technique has its own features Segmentation will be logical implementation in our research, which is generally used to find the region of interest on the behalf of some characteristic of the images. Initially, dilation is applied then ROI is calculated. The suspicious region with possibilities of kidney stone has been determined by using the characteristics and features of images.

### 3.5 *Region of Interest (ROI)*

The region of interest is a part of an image on which filtering or various other operations can be implemented. The ROI can be created in many shapes, to use the high-level ROI functions, such as draw circle or draw polygon. There may be more than one ROI in an image. By using ROI, images have a binary mask that implies the pixels belonging to ROI are denoted by 1 otherwise 0. To identify the kidney stone, kidney region is focused. We use the range criteria to select the label that denotes the stone region and to exclude the unlikely labels (tape artifacts and high-intensity labels). Finally, to obtain the effective kidney stone region, the result of this step was multiplied with the original image. Thus kidney stone is detected [20]. The result of renal calculi images is used for future analysis. From the renal calculi image, the calculi regions are extracted. For ROI creation classes used are—`images.roi.Circle` or `image.roi.Polygon`.

### 3.6 *Feature Extraction*

The quality of ultrasound medical images is tested based on various texture features. The main target of the features extraction is to collect essential features of the region to be investigated in the kidney image [2]. The features are being extracted to calculate the different parameters like contrast, entropy, energy, and correlation from the kidney stone image (Table 1).

## 4 **Confusion Matrix**

It comprises of real and expected classifications information carried out by a KNN classifier [1]. The efficacy of the proposed region of interest detection system is evaluated by analysis of the confusion matrix.

Where the number of actual negative cases in the data = Condition Negative (N)

The number of actual positive cases in the data = Condition Positive (P)

No. of correct positive prediction = True Positive (TP)

No. of correct negative Prediction = True Negative (TN)

**Table 1** GLCM features and their description

S. No.	Feature	Description
1	Contrast	It expresses the dissimilarity among the darkest and lightest regions in an image $\text{Contrast} = \sum_{i,j} (i - j)^2 X(i, j)$ where, $i$ and $j$ are the pixel values
2	ASM or Energy (Angular second moment)	It is the quality representing homogeneity of an image. ASM higher value shows textural uniformity $\text{ASM} = \sum_{i,j=0}^{Ng-1} p(i, j)^2$ where, $Ng$ is gray tone image
3	Entropy	The entropy also known as intensity distribution. This measures the randomness in the texture of an image $\text{Entropy} = - \sum_{i,j=0}^{Ng-1} p(i, j) \log(p(i, j))$ The lower entropy value indicates more homogeneity of an image and vice versa
4	Correlation	The Correlation performance evaluating parameter measures the relationship between two variables and mathematically given by-Correlation = $\frac{Cov(x,y)}{\sigma_x \sigma_y}$

No. of incorrect positive prediction, type I error = False Positive (FP)

No. of incorrect negative predictions, type II error = False Negative (FN)

Sensitivity or true positive rate (TPR)

$$\text{TPR} = \frac{\text{TP}}{\text{P}} = \frac{\text{TP}}{\text{TP} + \text{FN}} = 1 - \text{FNR} \tag{3}$$

Specificity or true negative rate (TNR)

$$\text{TNR} = \frac{\text{TN}}{\text{P}} = \frac{\text{TN}}{\text{TN} + \text{FP}} = 1 - \text{FPR} \tag{4}$$

Precision or positive predictive value (PPV)

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{5}$$

**Accuracy:** The accuracy of the classification process is based on correct and incorrect predictions. Following formula used to calculate the accuracy of the classification process [2].



Accuracy (ACC)

$$ACC = \frac{TP + TN}{P + N} = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

F1 SCORE is the harmonic mean of precision and sensitivity

$$F1 \text{ SCORE} = 2 \cdot \frac{PPV \cdot TPR}{PPV + TPR} = \frac{2TP}{2TP + FP + FN} \quad (7)$$

## 5 Result and Analysis

In this work, out of a total of 30 images, 15 with stone and 15 without stone are being considered. The statistical analysis of the presented technique is carried out on a set of kidney ultrasound images. For detecting the kidney stones, the segmentation followed by classification techniques has been used. In initial step of implementation, the dataset has been taken in the form of ultrasound image as shown in Fig. 6 [25]. After reading the images in Matlab, all related labels or text inside the given images has been removed. Further, the images are converted from RGB to Gray as shown in Fig. 7. In Fig. 7 the initial image was colored, but after conversion process it was converted to Gray Image.

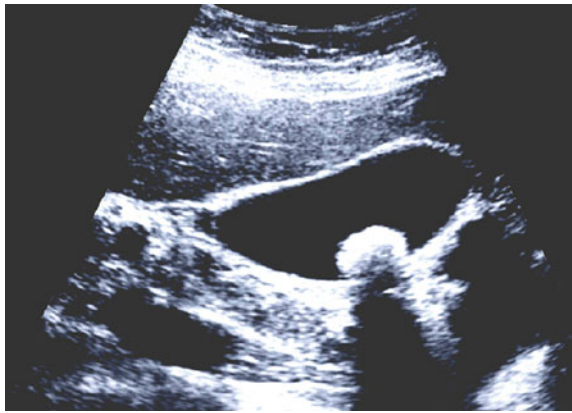
The visibility of low contrast ultrasound images has been enhanced by using contrast adjustment through MATLAB simulation. The enhanced image is captured in Fig. 8, as low contrast image has low visibility but after enhancing it becomes more clear. Further, the speckle noise has been eliminated by using median filter. This makes the image more visible as depicted in Fig. 9. Finally, to find the kidney stone or region of interest, morphological segmentation operations are utilized. The dilation operation is applied then the ROI is calculated as shown by Figs. 10 and 11.

**Fig. 6** Input image (kidney with stone)



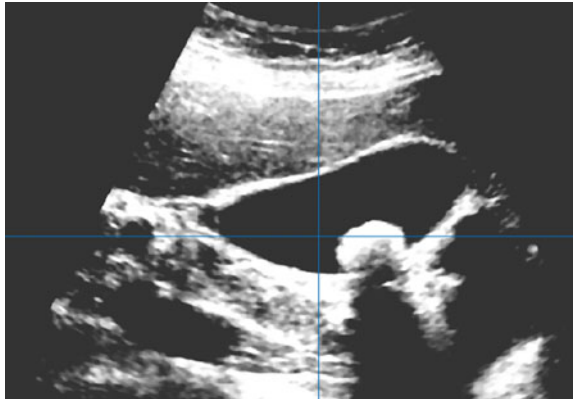


**Fig. 7** Preprocessed image



**Fig. 8** Enhanced image

**Fig. 9** Filtered image



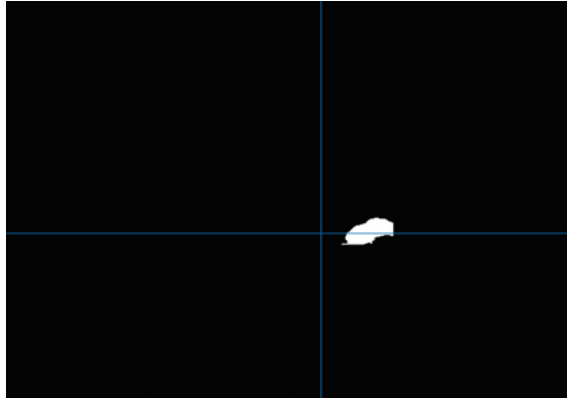


Fig. 10 Morphological segmented image

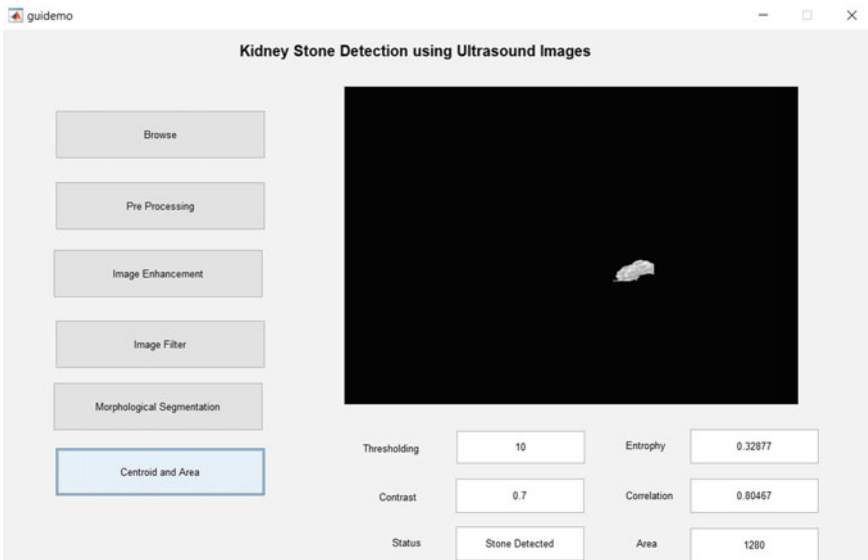


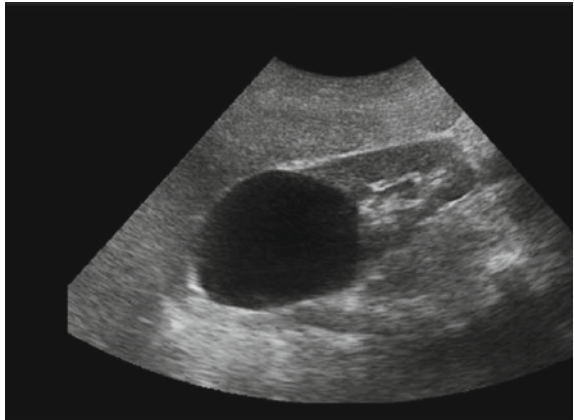
Fig. 11 Gui for kidney stone detection (centroid and area)

Figure 10 shows the final segmented image. The various parameters such as contrast, entropy, Threshold, correlation and area of stone are shown in Fig. 11. The Figs. 12, 13, 14, 15, 16 and 17 are preprocessed, enhanced, filtered, and segmented as Figs. 6, 7, 8, 9, 10 and 11. These images have no stones. Hence the ROI obtained using Matlab is blank as shown in Fig. 17. Figure 18 shows the performance of final system at threshold 10, 15 and 20. The Table 2 shows the confusion matrix for a two-class classifier. For Analysis Confusion matrix is used. It helps in detecting the accuracy, F1

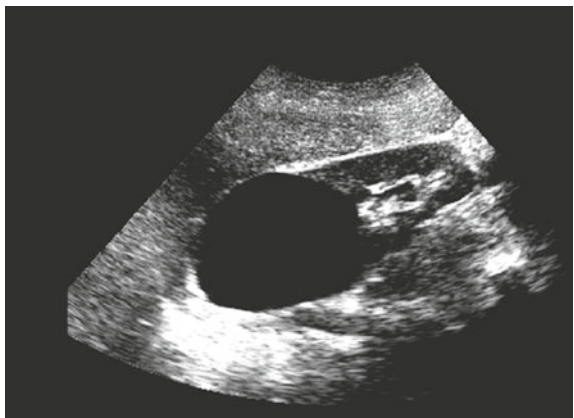
**Fig. 12** Input image (kidney without stone)

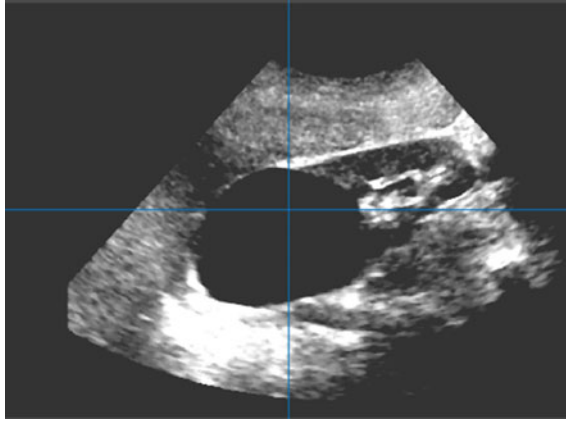


**Fig. 13** Preprocessed image

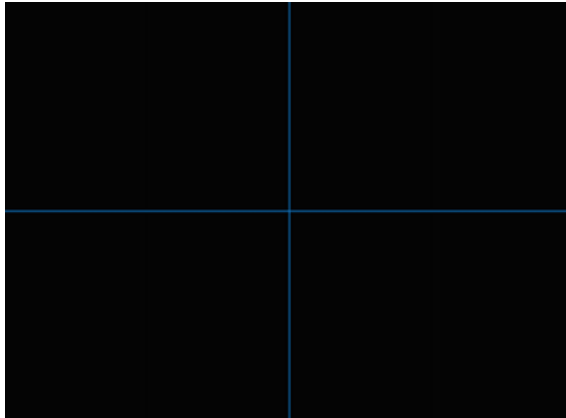


**Fig. 14** Enhanced image





**Fig. 15** Filtered image



**Fig. 16** Morphological and segmented image

Score, sensitivity and specificity of datasets at 10, 15, and 20 segmentation threshold as shown in Table 5 (Figs. 19 and 20).

The confusion matrix for a two-class classifier, The confusion matrix is shown in Table 2.

The threshold segmentation is varied as 10, 15, and 20. This has been observed that beyond 20 the images are not visible (Tables 3 and 4).

Table 5 shows that the best accuracy is obtained at a threshold value of 20.

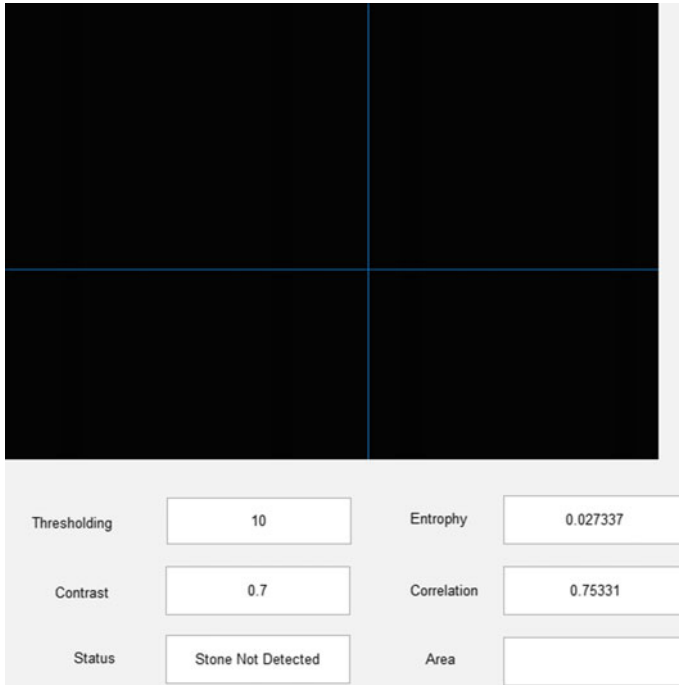


Fig. 17 GUI for kidney stone detection (centroid and area)

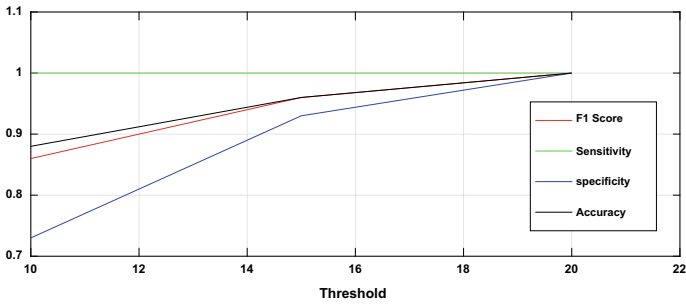


Fig. 18 Performance evaluation of final system

Table 2 Confusion matrix

		Predicted	
		Negative	Positive
Actual	Negative	TN	FN
	Positive	FP	TP

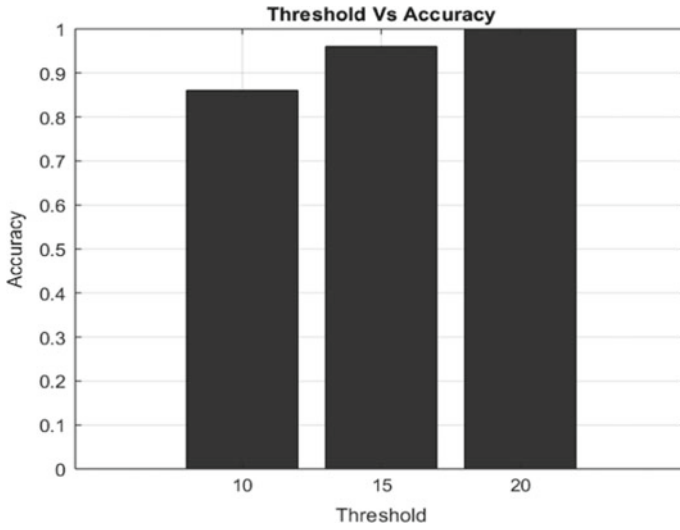


Fig. 19 Threshold versus accuracy

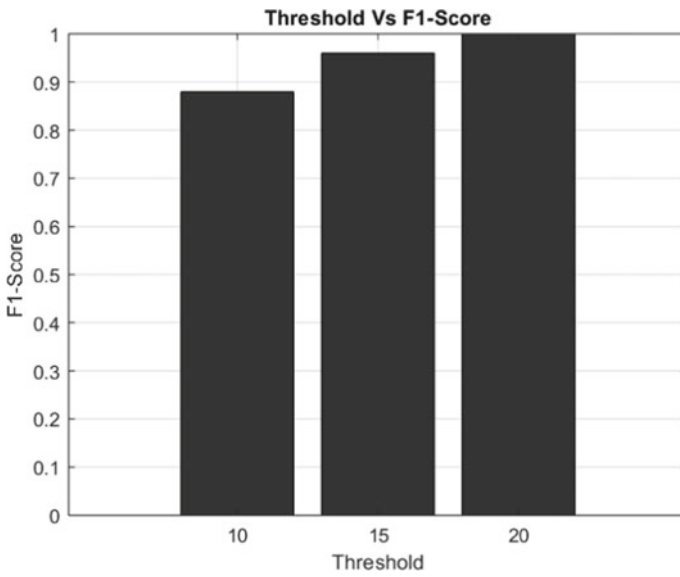


Fig. 20 Threshold versus F1 score

**Table 3** Segmentation thresholding as ( $\delta = 10$ )

S. No	True positive (TP)	True negative (TN)	False positive (FP)	False negative (FN)
1	1	0	0	0
2	1	0	0	0
3	1	0	0	0
4	1	0	0	0
5	1	0	0	0
6	1	0	0	0
7	1	0	0	0
8	1	0	0	0
9	1	0	0	0
10	1	0	0	0
11	1	0	0	0
12	1	0	0	0
13	1	0	0	0
14	1	0	0	0
15	1	0	0	0
16	0	1	0	0
17	0	0	1	0
18	0	0	1	0
19	0	0	1	0
20	0	1	0	0
21	0	1	0	0
22	0	1	0	0
23	0	1	0	0
24	0	1	0	0
25	0	1	0	0
26	0	1	0	0
27	0	1	0	0
28	0	0	1	0
29	0	1	0	0
30	0	1	0	0
Total	15	11	4	0

**Table 4** Segmentation thresholding as ( $\delta = 10$ )

Sensitivity	1
Specificity	0.733333333
Precision	0.789473684
Accuracy	0.866666667
F1-score	0.882352941



**Table 5** Simulation results for various threshold segmentation

S. No	Threshold	Accuracy	F1 score	Sensitivity	Specificity
1	10	0.86	0.88	1	0.73
2	15	0.96	0.96	1	0.93
3	20	1	1	1	1

## 6 Conclusion

The ultrasound images are not clear many times and require an additional diagnosis. In this work the early and accurate detection of kidney stones is proposed using ultrasound images. The proposed technique involves preprocessing, enhancement, filtration followed by morphological segmentation. The comparative analysis using GLCM features and confusion matrix verifies the efficacy of the proposed technique. The technique may be highly helpful for medical practitioners for accurate and early detection of kidney stones.

## References

1. Ebrahimi, S., & Mariano, V. Y. (2015). Image quality improvement in kidney stone detection on computed tomography images. *Journal of Image and Graphics*, 3(1).
2. Monika, P., Harsh, S., Sukhdev, S. (2016). A technique to suppress speckle in ultrasound images using nonlocal mean and cellular automata. *Indian Journal of Science and Technology*, 9(13).
3. Vasanthselvakumar, R., Balasubramanian, M., & Palanivel, S. (2017). Pattern analysis of kidney diseases for detection and classification using ultrasound b-mode images. *International Journal of Pure and Applied Mathematics*, 117(15), 635–653.
4. Sharma, K., & Jitendra, V. (2017). A decision support system for classification of normal and medical renal disease using ultrasound images: a decision support system for medical renal diseases. *International Journal of Ambient Computing and Intelligence (IJACI)*, 8(2), 52–69.
5. Torres, H. R., et al. (2018). Kidney segmentation in ultrasound, magnetic resonance and computed tomography images: A systematic review. *Computer Methods and Programs in Biomedicine*, 157, 49–67.
6. Saroha, V., Verma, R., & Saini, K. K. (2016). Enhancement of kidney stone images using different filtering technique. *International Journal of Scientific Research and Management*, 4(8).
7. Selvarani, S., & Rajendiran, P. (2018). Feature fusion based hybrid classifier to detect the renal abnormalities in ultrasound imaging.
8. Bharti, P., Mittal, D., & Ananthasivan, R. (2018). Preliminary study of chronic liver classification on ultrasound images using an ensemble model. *Ultrasonic Imaging*, 40(6), 357–379.
9. Kaur, R. (2017). Review of image segmentation technique. *International Journal of Advanced Research in Computer Science*, 8(4).
10. Verma, J., et al. (2017). Analysis and identification of kidney stone using Kth nearest neighbour (KNN) and support vector machine (SVM) classification techniques. *Pattern Recognition and Image Analysis*, 27(3), 574–580.
11. Vaish, P., et al. (2016). Smartphone based automatic abnormality detection of kidney in ultrasound images. In *2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom)*. IEEE.

12. Nithyavathy, N., et al. (2016). Ultrasound imaging technique for the identification of kidney stones using Gsd platform. *Ultrasound*, 5(1).
13. Hu, Z., & Tang, J. (2016). Cluster driven anisotropic diffusion for speckle reduction in ultrasound images. In *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE.
14. Viswanath, K., & Ramalingam, G. (2014). Design and analysis performance of kidney stone detection from ultrasound image by level set segmentation and ANN classification. In *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. IEEE.
15. Rahman, T., & Uddin, M. S. (2013). Speckle noise reduction and segmentation of kidney regions from ultrasound image. In *2013 International Conference on Informatics, Electronics & Vision (ICIEV)*. IEEE.
16. Tijjani, A., Hashim, U., & Sani, U. S. (2012). Designing an artificial neural network model for the prediction of kidney problems symptom through patient's metal behavior for pre-clinical medical diagnostic. In *2012 International Conference on Biomedical Engineering (ICoBE)*. IEEE.
17. Jeyalakshmi, T. R., & Ramar, K. (2010). A modified method for speckle noise removal in ultrasound medical images. *International Journal of Computer and Electrical Engineering*, 2(1), 54.
18. Hafizah, W. M., Supriyanto, E., & Yunus, J. (2012). Feature extraction of kidney ultrasound images based on intensity histogram and gray level co-occurrence matrix. In *2012 Sixth Asia Modelling Symposium (AMS)*. IEEE.
19. Bharath, R., et al. (2015). FPGA-based portable ultrasound scanning system with automatic kidney detection. *Journal of Imaging*, 1(1), 193–219.
20. Bora, D. J. (2017). Importance of image enhancement techniques in color image segmentation: A comprehensive and comparative study. arXiv preprint [arXiv:1708.05081](https://arxiv.org/abs/1708.05081).
21. Goel, R., & Jain, A. (2018). The implementation of image enhancement techniques on color n gray scale IMAGES. In *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)*.
22. Raja, R. A., & Jennifer Ranjani, J. (2013). Segment based detection and quantification of kidney stones and its symmetric analysis using texture properties based on logical operators with ultra sound scanning. In *International Conference on Computing and Information Technology (ICIT'2013)*.
23. Meiburger, K. M., Rajendra Acharya, U., & Molinari, F. (2018). Automated localization and segmentation techniques for B-mode ultrasound images: A review. *Computers in Biology and Medicine*, 92, 210–235.
24. Zheng, Q., et al. (2018). A dynamic graph cuts method with integrated multiple feature maps for segmenting kidneys in 2D ultrasound images. *Academic Radiology*.
25. Monika, P., Harsh, S., & Sukhdev, S. (2016). Features extraction and classification for detection of kidney stone region in ultrasound images. *International Journal of Multidisciplinary Research and Development*, 3(5), 81–83.

# Non-adaptive and Adaptive Filtering Techniques for Fingerprint Pores Extraction



Diwakar Agarwal and Atul Bansal

**Abstract** Fingerprint sweat pores as Level 3 features have the capability to improve the accuracy of the fingerprint recognition process. Extraction of pores is the foremost step in the designing of fingerprint high-level features based applications such as migration control at the borders, identification of fake fingerprints, etc. Filtering based approach is one of the prominent techniques for pores extraction. All the available filtering based methods are categorized into non-adaptive and adaptive techniques. This paper presents an extended description of both the techniques. Some practical concerns while implementing the methods have also been highlighted. Experimental results have been carried out on the high resolution database of 500 dpi fingerprint images. Performance has been measured and compared on the basis of true detection rate ( $R_T$ ) and false detection rate ( $R_F$ ). Simulation results show that the adaptive filtering techniques achieve better  $R_T$  and  $R_F$ .

**Keywords** Adaptive · Filtering · Biometrics · Detection accuracy · Fingerprint · Non-adaptive filtering · Sweat pores

## 1 Introduction

Biometrics is aiming to study and design the technical solutions to evaluate quantitatively the human characteristics (both behavioral and physical) for its categorization. Among all the biometrics (face, palmprint, fingerprint, speech, iris, signature, gait, ear, etc.), fingerprint is one of the most mature and proven technology. Fingerprint ridge details are generally categorized hierarchically into three levels feature. Level 1 includes extensive details of fingerprint such as ridge flow and patterns [1]. Level 2 features comprised of minutia points (ridge bifurcation and ending) and Level 3

---

D. Agarwal (✉) · A. Bansal

Electronics & Communication Engineering, GLA University, 17 Km stone, NH-2, Mathura-Delhi Road, Mathura, UP, India

e-mail: [diwakar.agarwal@gla.ac.in](mailto:diwakar.agarwal@gla.ac.in)

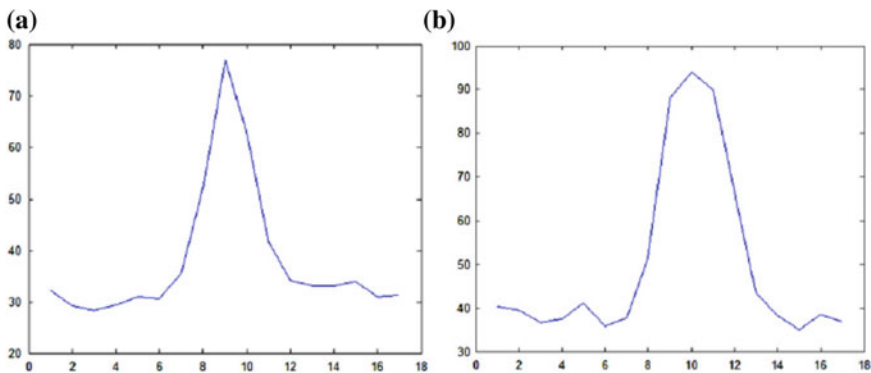
A. Bansal

e-mail: [atul.bansal@gla.ac.in](mailto:atul.bansal@gla.ac.in)

covers sweat pores and ridge dimensional properties such as ridge width, deviation, boundary contours, etc. [2]. Traditional Automated Fingerprint Identification System (AFIS) which depends upon level 1 and level 2 features are vulnerable to presentation attacks by the hackers. Therefore, the successful implementation of the biometric system relies upon the Level-3 features. Recently, fingerprint sweat pores as level 3 features are utilized for the purpose of identification [2–5] and also for the prevention of spoofing [6–12]. Pores are unique, varying in size and evenly distributed along the ridges. Pores may be open or closed depending upon the perspiration activity [13].

Pore detection methods are generally classified into skeletonization based approach and filtering based approach. Earlier work by Stosz and Alyea [3] and Kryszczuk et al. [14] are based on skeletonization. Fingerprint ridge patterns are determined simply by tracing the line segments while maintaining the continuity of the pixels. The disadvantages of this approach is three folds: (a) Skeletonization only works on very high-quality images (b) Alignment of test and query region is required for fingerprint matching (c) It does not provide reliable results when skin conditions are not favorable in case of wet skin, dry skin, scars, wrinkles, etc. In contrast to the skeletonization, filtering based approaches [2, 4, 15–18] are based on the 2D convolution of the input fingerprint image and the pore model (acting as a filter). This kind of approach is based upon the concept of matched filter where the output of filtering hits maxima or minima (depends on whether dark object surrounded by bright background or bright object over a dark background, respectively) whenever the pore model under consideration exactly lay upon the pore. The real pore on the fingerprint has Gaussian shaped intensity profile as shown in Fig. 1 [4]. The most important parameter in 2D Gaussian function is its scale value, i.e., standard deviation ( $\sigma$ ) which measures the spreadness in 2D space. In order to get the highest matched filter response, it is required that the dimension of the considered Gaussian pore model coincide with the pore size underneath it.

The contribution of the work is as follows: (a) Two Non-adaptive [2, 15] and one adaptive [4] filtering methods have been implemented. Some practical issues including advantages and drawbacks have also been investigated while implementing



**Fig. 1** Intensity profile across **a** closed pore, **b** open pore [4]

the algorithms. (b) All the three methods have been applied on the LivDet 2013 Biometrika test database [19] and compared on the grounds of  $R_T$  and  $R_F$ . This paper is organized as follows: Sect. 2 provides the literature review of the various filtering based pore detection methods. Sections 3 and 4 includes the description and results of the implementation of non-adaptive and adaptive techniques, respectively. The conclusion is provided in Sect. 5.

## 2 Existing Filtering Based Pore Detection Methods

Pore detection methods are classified as non-adaptive and adaptive filtering based technique. Some literature is available on non-adaptive technique of pore detection [2, 15–17]. In 2005, Ray et al. [15] were first introduced the pore model which is equivalent to the modified 2D Gaussian function. Here, high and low probable pore areas are first identified by taking the sum of squared error in  $3 \times 3$  local neighborhood. The Gaussian function has fixed scale value hence; Ray's pore model is unitary. It is also isotropic since appearance of real pores is considered constant for all the regions of the fingerprint. In 2007, Jain et al. [2] utilized the Mexican hat wavelet to capture the high negative frequency response which is generated by the low intensity blobs. The scale value of the pore model is considered constant for all the pores of the fingerprint. Hence, the model is non-adaptive to the various sizes of the real pore encountered in the whole fingerprint. In 2008, Parsons et al. [16] introduced DoG (Difference of Gaussian), based pore model. Here, the edges of the pores are determined by using fixed scale value of the Gaussian function. In 2010, Manivanan et al. [17] proposed highpass and correlation filtering for the detection and localization of active sweat pores, respectively.

There are some literature which are available on adaptive technique [4, 18]. In 2010, Zhao et al. [18] proposed an adaptive DoG method. The block-wise approach is implemented and the scale value is adaptively determined in relation to local ridge frequency. The shape of the pores is assumed as circular which is not valid for the real sweat pores. Hence, the pore model is isotropic. Zhao et al. [4] proposed an adaptive anisotropic pore model to overcome the problem of a fixed scale that has been prevalent in [2, 15–17]. The scale value and the orientation of the model are adaptively determined by local ridge frequency and local ridge orientation, respectively. Moreover, several literature on pore detection are reported based on Convolutional Neural Network (CNN) [20–24]. Jang et al. [20] utilized ten learning layers together with ReLu for the pore detection & intensity refinement. Labati et al. [21] proposed a method for heterogeneous dataset which includes touch-based, touchless & latent fingerprints. Su et al. [22] proposed the method for arbitrary size images. Dahia et al. [23] describe the supervised method for automatic dataset annotations. Genovese et al. [24] also report a CNN based method for pore detection in touchless fingerprints.

### 3 Non-adaptive Pore Detection Technique

In non-adaptive filtering technique, the scale of the pore model is fixed and remains constant for all the pores of the fingerprint. This is a kind of isotropic filtering since empirically chosen unitary scale parameter is considered throughout the image. In this section, results of pore detection methods implemented by Ray et al. [15] and Jain et al. [2] are shown. Some practical issues in the implementation of these methods have also been discussed.

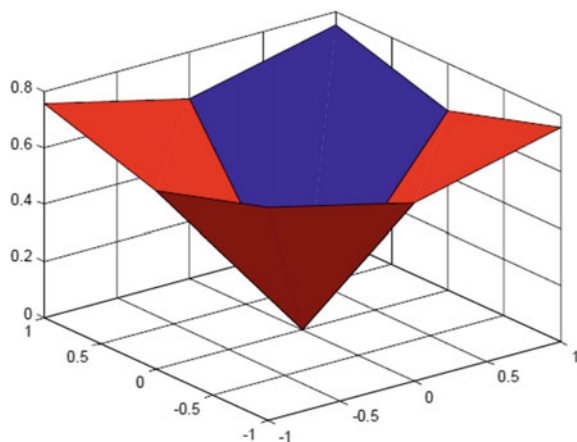
#### 3.1 Pore Detection Method Implemented by Ray et al. [15]

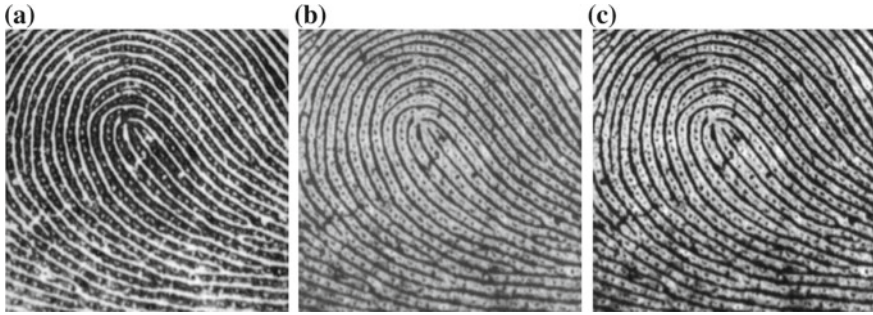
In 2005, Ray et al. [15] proposed a pore detection method for 500 dpi resolution fingerprint images. This method utilizes a modified 2D Gaussian function as a pore model. The mathematical expression of the pore model under consideration is given by (1) and its 3D shaded surface representation is shown in Fig. 2.

$$G(x, y) = 1 - e^{-|x^2+y^2|^{\frac{1}{2}}} \quad (1)$$

For the purpose of the implementation, one fingerprint image is taken from the database as an example image as shown in the Fig. 3a. The process begins by performing the image negative transformation (2) on the original gray scale image  $I(x, y)$ . The output of the transformation is shown in the Fig. 3b in which ridges and valleys are interpreted by bright and dark region respectively. Next, the image normalization (3) has been performed so that the intensity values exist between 0 and 1 as shown in Fig. 3c.

**Fig. 2** 3D shaded surface representation of modified 2D Gaussian function





**Fig. 3** **a** Original gray scale fingerprint image, **b** image negative image, **c** normalized image

$$N(x, y) = 255 - I(x, y) \quad (2)$$

$$F(x, y) = \frac{N(x, y)}{255} \quad (3)$$

Error map has been determined by using (4) in order to measure the similarity between pore model and the real pore in the fingerprint image.

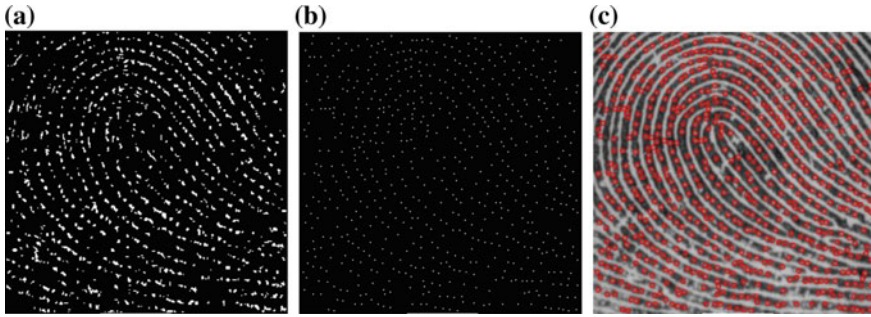
$$E(x, y) = \sum_{i=x-r}^{x+r} \sum_{j=y-r}^{y+r} (F(i, j) - G(i - x + r, j - y + r))^2 \quad (4)$$

where  $G$  is the pore model and  $r$  is the distance from the central pixel to the boundary of the pore model which is equal to 1 for  $3 \times 3$  size pore model. The dimension of the pore model remains same for all the pores of the fingerprint. This shows the isotropic nature of the algorithm. Error map exposes those areas in fingerprint image which have high possibility of located pores. Exactly matched pore model with the pore region generates low error which tends to the high possibility of the presence of pores. Thus, a thresholding operation has been performed over error map in order to generate binarized error map  $E_B(x, y)$  (see Fig. 4a) where binary '1' represents low error region. To find out the accurate location of sweat pores, error points in  $E(x, y)$  that own minimum error in a  $7 \times 7$  local neighborhood  $E_L$  have been extracted. Later, these points are represented as binary '1' by using (5). Therefore, a binary image of local minimum value error points  $B_L(x, y)$  is formed which is shown in the Fig. 4b.

$$B_L(x, y) = \begin{cases} 1, & E(x, y) = \min[E_L] \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

where the local neighborhood  $E_L$  is defined by (6)

$$E_L = E(x - r_m : x + r_m, y - r_m : y + r_m) \quad (6)$$



**Fig. 4** **a** Binarized error map **b** Binary image of local minimum value error points **c** Pores highlighted by red circle. True detection rate ( $R_T$ ) = 81.3% and false detection rate ( $R_F$ ) = 22.2%

The binary image  $B_L(x, y)$  has been filtered by the binarized error map  $E_B(x, y)$  in order to remove any spurious error point which may arise from non-pore regions. Thus final binary pore map  $P(x, y)$  has been obtained by performing the simple masking operation as shown in (7). Figure 4c shows the pores valid locations which are highlighted on the original fingerprint image. Equations (8) and (9) define the computation of the values of  $R_T$  and  $R_F$ .

$$P(x, y) = E_B(x, y)B_L(x, y) \quad (7)$$

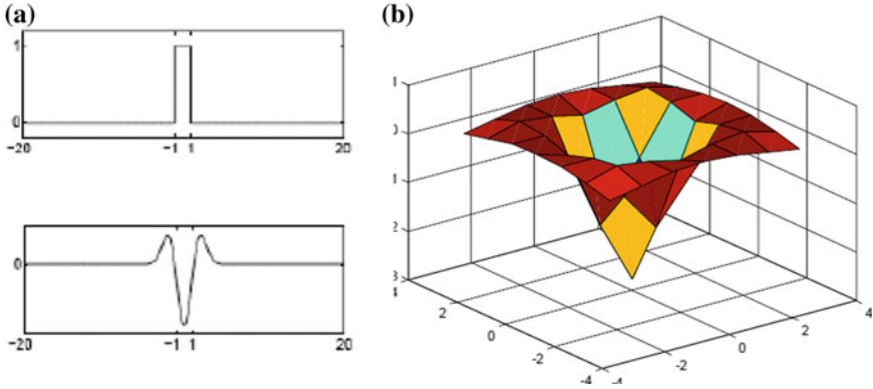
$$R_T = \frac{\text{Number of detected true pores}}{\text{Total Number of all true pores}} \quad (8)$$

$$R_F = \frac{\text{Number of false detected pores}}{\text{Total number of all detected pores}} \quad (9)$$

### (1) Practical implementation issues

- (a) In this method, the dimension of the pore model is fixed and empirically chosen as  $3 \times 3$ . The pore model larger than  $3 \times 3$  does not fit into real pore in the fingerprint image. Therefore, high error points will be generated in the error map which are later wiped away from the binarized error map after thresholding. As a consequence, less number of probable pore points appeared in the final pore map.
- (b) The threshold value is empirically chosen as 0.5. If the threshold value is chosen larger than 0.5, there will be the chances that some non-pore regions are falsely considered as true pore regions (cloud like regions in binarized error map). This will generate spurious pores later in the binary image of local minimum value error points. Whereas, the threshold value smaller than 0.5 leads to the absence of actual pore points.
- (c) The dimension of local neighborhood  $E_L$  is also fixed and empirically chosen as  $7 \times 7$ . The appropriate size of  $E_L$  is necessary to control the





**Fig. 5** a Abrupt intensity changes at the edges, b 3D shaded surface representation of the pore model

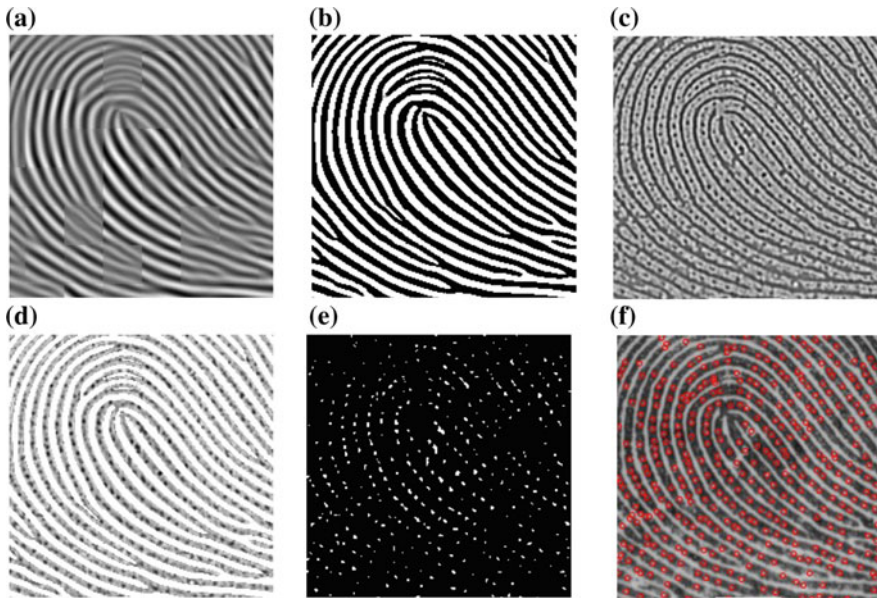
- number of multiple pore locations which are appeared on the same real pore.
- (d) In this method, there is no provision to eliminate false pore locations which are appeared on the valleys due to some artifacts.

### 3.2 Pore Detection Method Implemented by Jain et al. [2]

In 2007, Jain et al. [2] imposed another filtering method for the extraction of sweat pores. Pores are observed as a bright object over dark ridges. Therefore, the intensity changes abruptly from dark to bright and from bright to dark at the edges. An analogous example of abrupt intensity change is shown in Fig. 5a. This method utilized the Mexican hat wavelet as a pore model in order to capture high negative frequency response. The mathematical expression of the pore model under consideration is given by (10) and its 3D shaded surface representation is shown in Fig. 5b.

$$w(s, a, b) = \frac{1}{\sqrt{s}} \iint f(x, y) \varphi\left(\frac{x-a}{s}, \frac{y-b}{s}\right) dx dy \tag{10}$$

Jain et al. [2] utilized the fact that sweat pores follow the structure of the ridges and being determined as long as the friction ridges are identified. Therefore, at first, fingerprint ridge enhancement has been done and then pores are extracted along the ridges. In order to enhance the ridges, the method implemented by Hong et al. [25] is utilized. Fingerprint ridge enhanced image is shown in Fig. 6a and the binary image in which ridges are represented in black and valleys in white is shown in Fig. 6b. After filtering with the pore model, pores appear as low-intensity small blobs in the output image as shown in Fig. 6c. Then, the linear combination of the ridge enhanced image



**Fig. 6** **a** Fingerprint ridge enhanced image, **b** ridge enhanced binary image, **c** result after 2D filtering, **d** linear combination of ridge enhanced image and output filtered image, **e** binarized pore map, **f** pores highlighted by red circle.  $R_T = 87.5\%$  and  $R_F = 18.1\%$

and output filtered image has been done for the purpose of the optimal enhancement of pores along the ridges as shown in Fig. 6d. The Global thresholding is performed on Fig. 6d) in order to segment out the probable pore points. The threshold value is empirically chosen as 125 below which all the pixels are represented by binary '1'. Then, the segmented blobs which are larger than 40 pixels are eliminated. The final binarized pore map is shown in Fig. 6e. Figure 6f shows the pore valid locations which are highlighted on the input fingerprint image.

#### (1) Practical implementation issues

- (a) The scale parameter of the Mexican hat wavelet defines the depth of the filter. The larger the scale value less will be the height of the negative lobe of the pore model. As a consequence, less negative responses which are generated by pores will be captured. Whereas, for the smaller scale values, deep negative lobe of the pore model will detect even the slight change in intensity that could possibly be generated through some artifacts. This will lead to the generation of spurious pores. Hence, the trade-off lies in choosing the scale value. This clearly shows the isotropic nature of the pore model as the same scale value is used for all the pores of the fingerprint.
- (b) The global thresholding is performed over the whole image. The high threshold value will consider some non-pore regions as probable pore candidate. This will result in false locations of pores in the binarized pore map.

- (c) The value of true detection rate is greater than the value obtained from Ray’s method. Correspondingly, the value of the false detection rate reached low value.

### 4 Adaptive Pore Detection Technique

In contrast to non-adaptive filtering, the scale factor in the adaptive technique varies with respect to the different regions of the fingerprint. The adaptive technique is the solution to the problem of the intra-region variation of the size of real pores. In 2010, Zhao et al. [4] introduced an adaptive and anisotropic pore model as defined by (11).

$$P_o(i, j) = e^{-\frac{j^2}{2\sigma^2}} \cos\left(\frac{\pi}{3\sigma}i\right), -3\sigma \leq i, j \leq 3\sigma \tag{11}$$

where ‘ $\sigma$ ’ is the scale parameter. In a local region, fingerprint ridges are run approximately parallel to each other and oriented by the common angle. Hence, each region has been designated by the local ridge frequency and the local ridge orientation which are determined by following Hong et al. [25] method.

Basically, the adaptiveness of the pore model relies upon region-wise determination of the orientation and the scale value of the pore model. In this method, the rotation angle of the pore model has been determined from the local ridge orientation. The rotated pore model is given by (12) and Fig. 7 shows two different orientations of the pore model as an example.

$$P_\theta(i, j) = \text{Rot}(P_o, \theta) \tag{12}$$

The scale of the pore model which is given by (13) has been determined from the local ridge frequency.

$$\sigma = \frac{\tau}{k} \tag{13}$$

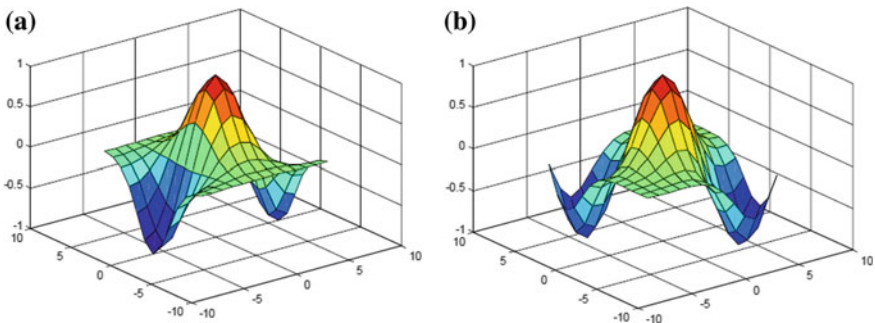
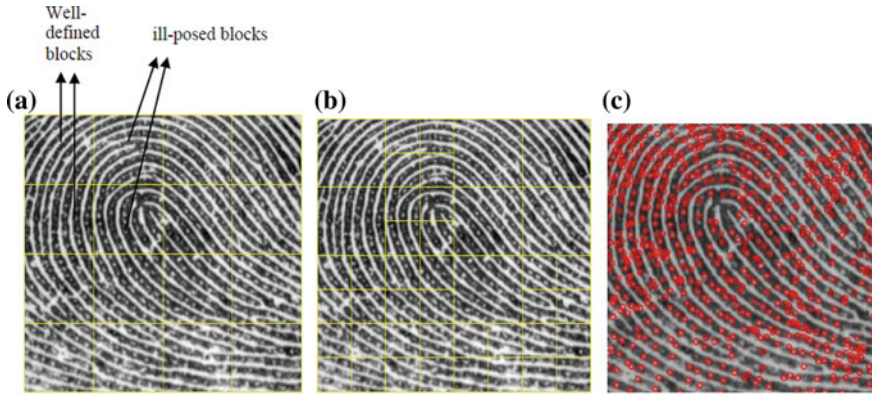


Fig. 7 Pore models **a** Basic pore model at  $\theta = 0$ , **b** rotated pore model at  $\theta = 45$



**Fig. 8** **a** Categorization of blocks into well-defined and ill-posed blocks **b** partitioning of ill-posed blocks into four equal sized sub blocks **c** Final pore map.  $R_T = 88.2\%$  and  $R_F = 18.0\%$

Here,  $\tau = \frac{1}{f}$  is the local ridge period and ‘ $k$ ’ is the positive constant. There are some regions in the fingerprint where ridge orientation and frequency cannot be determined properly due to some irregularities or due to some singular points present in the fingerprint. Therefore, in this method, a block-wise approach is proposed. The whole fingerprint image is divided into multiple blocks of fixed size and categorized them as well-defined and ill-posed blocks as shown in Fig. 8a. All ill-posed blocks are further divided into four equal sized sub blocks (see Fig. 8b) and categorized them again. Well-defined blocks are equipped with both dominant ridge orientation and frequency thus filtered by adaptive pore model. Ill-posed blocks do not have dominant frequency hence their frequency is determined through the interpolation of the frequencies of neighboring well-defined blocks. Ill-posed blocks are filtered by adaptive DoG pore model which was proposed by Zhao et al. [18]. Then, thresholding followed by post-processing has been performed in order to generate final pore map as shown in Fig. 8c.

#### (1) Practical implementation issues

- (a) In this method, the scale value of the pore model is adaptively determined for all the regions of the fingerprint through the block-wise approach. As a result, the possibility of the detection of real pores in a local region will be high as compared to non-adaptive methods.
- (b) Some true pores may be unaccounted since global thresholding is used to segment out the candidate pore region.
- (c) The value of true detection rate reached a high value in comparison to Ray’s and Jain’s method whereas; false detection rate reached low value.

Performance comparison of all three pore detection methods is shown in Table 1.

**Table 1** Performance comparison of non-adaptive and adaptive filtering based pore detection methods

Pore detection method	Technique	$R_T$ (%)	$R_F$ (%)
Ray et al. [15]	Non-adaptive	81.3	22.2
Jain et al. [2]	Non-adaptive	87.5	18.1
Zhao et al. [4]	Adaptive	88.2	18.0

## 5 Conclusion

The rapid deployment of high-level features based technology in biometric systems has completely changed the current automated fingerprint recognition. Sweat pore is one of the most quantifying high-level features used for various applications like border security control, fingerprint liveness detection, migration control, etc. It is concluded that adaptive filtering based pore detection methods are designed according to the anisotropic nature of the real pore. Hence, less number of false pores appears in comparison to the non-adaptive filtering methods. As a future aspect, researchers can integrate CNN and adaptive nature of pores for the designing of complete automatic pore detection method.

**Acknowledgements** Authors would like to thank the principal investigator of the Department of Electrical & Electronic Engineering, University of Cagliari. Permission has been granted and authentic credentials have been issued by the University in order to Access the LivDet Joint Multimodal Biometric Dataset 2013.

## References

- Galton, F. (1889). Personal identification and description. *Journal of Anthropological Institute of Great Britain and Ireland*, 177–191.
- Jain, A. K., Chen, Y., & Demirkus, M. (2007). Pores and ridges: Fingerprint matching using level 3 feature. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(1), 15–27.
- Stosz, J., & Alyea, L. (1994). Automated system for fingerprint authentication using pores and ridge structure. *Proceedings of SPIE*, 22(77), 210–223.
- Zhao, Q., Zhang, D., Zhang, L., & Luo, N. (2010). Adaptive fingerprint pore modeling and extraction. *Pattern Recognition*, 43(8), 2833–2844.
- Ashbaugh, D. R. (1999). *Quantitative-qualitative friction ridge analysis: An introduction to basic and advanced ridgeology*. CRC Press.
- Espinoza, M., & Champod, C. (2011). *Using the number of pores on fingerprint images to detect spoofing attacks* (pp. 1–5). Hong Kong: International Conference on Hand-Based Biometrics ICH.
- Johnson, P., Schuckers, S. (2014). Fingerprint pore characteristics for liveness detection. In *Proceedings IEEE BIOSIG*.
- Lu, M. Y., Chen, Z. Q., & Sheng, W. G. (2015, November). A pore-based method for fingerprint liveness detection. In *2015 International Conference on Computer Science and Applications (CSA)* (pp. 77–81). IEEE.
- Marcialis, G. L., Roli, F., & Tidu, A. (2010, August). Analysis of fingerprint pores for vitality detection. In *2010 20th international conference on pattern recognition* (pp. 1289–1292). IEEE.

10. Memon, S., Manivannan, N., & Balachandran, W. (2011, November). Active pore detection for liveness in fingerprint identification system. In *2011 19th Telecommunications Forum (TELFOR) Proceedings of Papers* (pp. 619–622). IEEE.
11. Da Silva, M. V., Marana, A. N., & Paulino, A. A. (2015, April). On the importance of using high resolution images, third level features and sequence of images for fingerprint spoof detection. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1807–1811). IEEE.
12. Abhyankar, A., & Schuckers, S. (2010, September). Towards integrating level-3 Features with perspiration pattern for robust fingerprint recognition. In *2010 IEEE International Conference on Image Processing* (pp. 3085–3088). IEEE.
13. Locard, E. (1912). Les Pores et L'Identification Des Criminels. *Biologica: Revue Scientifique de Medicine*, 2, 357–365.
14. Kryszczuk, K., Drygajlo, A., & Morier, P. (2004). Extraction of level 2 and level 3 features for fragmentary fingerprints. Proc. Second COST Action 275 Workshop: 83–88.
15. Ray, M., Meenen, P., & Adhami, R. (2005). A novel approach to fingerprint pore extraction. In *Proceedings of the 37th South-eastern Symposium on System Theory* (pp. 282–286).
16. Parsons, N. R., Smith, J. Q., Thonnes, E., Wang, L., & Wilson, R. G. (2008). Rotationally invariant statistics for examining the evidence from the pores in fingerprints. *Law, Probability and Risk*, 7, 1–14.
17. Manivanan, N., Memon, S., & Balachandran, W. (2010). Automatic detection of active sweat pores of fingerprint using highpass and correlation filtering. *Electronics Letters*, 46, 1268–1269.
18. Zhao, Q., Zhang, D., Zhang, L., & Luo, N. (2010). High resolution partial fingerprint alignment using pore-valley descriptors. *Pattern Recognition*, 43(3), 1050–1061.
19. Ghiani, L., Yambay, D., Mura, V., Tocco, S., Marcialis, G. L., Roli, F., & Schuckers, S. (2013). LivDet 2013—Fingerprint liveness detection competition 2013. In *6th IAPR/IEEE International Conference on Biometrics* (pp. 4–7).
20. Jang, H., et al. (2017). Deep pore: Fingerprint pores extraction using deep convolution neural networks. *IEEE Signal Processing Letters*, 24(12), 1808–1812.
21. Labati, R. D., Genovese, A., Muñoz, E., Piuri, V., & Scotti, F. (2018). A novel pore extraction method for heterogeneous fingerprint images using convolutional neural networks. *Pattern Recognition Letters*, 113, 58–66.
22. Su, H. R., Chen, K. Y., Wong, W. J., & Lai, S. H. (2017). A deep learning approach towards pore extraction for high-resolution fingerprint recognition. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 2057–2061). IEEE.
23. Dahia, G., & Segundo, M. P. (2018). CNN-based pore detection and description for high-resolution fingerprint recognition. ArXiv preprint [arXiv:1809.10229](https://arxiv.org/abs/1809.10229).
24. Genovese, A., Munoz, E., Piuri, V., Scotti, F., & Sforza, G. (2016). Towards touchless pore fingerprint biometrics: A neural approach. In *2016 IEEE Congress on Evolutionary Computation (CEC)* (pp. 4265–4272). IEEE.
25. Hong, L., Wan, Y., & Jain, A. K. (1998). Fingerprint image enhancement: Algorithms and performance evaluation. *IEEE Transaction Pattern Analysis and Machine Intelligence*, 20(8), 777–789.

# Character and Mesh Optimization of Modern 3D Video Games



Ragib Hasan, Sumittra Chakraborti, Md. Zonieed Hossain, Taukir Ahamed, Md. Abdul Hamid and M. F. Mridha

**Abstract** 3D video game assets, small props to larger mesh in a video game, make a huge difference in performance of video games. Non-optimized game assets may not support its presentation in wide range of hardware. Many people want to play video game but cannot play because of their lower end hardware setup. So, what is the point of developing a video game if vast amount of people cannot enjoy it? In this paper, we have introduced a simple but an effective methodology to optimize 3D mesh to support in wide range of hardware, which can satisfy the need of gamers. We have shown that reducing poly count and removing unnecessary parts from meshes contribute significantly to optimal performance.

**Keywords** 3D video games · Mesh optimization · Character optimization · Optimizing game asset · Inexpensive hardware · Indie developer

---

R. Hasan (✉) · S. Chakraborti · Md. Zonieed Hossain · T. Ahamed · Md. Abdul Hamid · M. F. Mridha  
Department of Computer Science and Engineering, University of Asia Pacific, Dhaka 1215, Bangladesh

e-mail: [md.ragib.supanta@gmail.com](mailto:md.ragib.supanta@gmail.com)

S. Chakraborti

e-mail: [sumittrachakraborti@gmail.com](mailto:sumittrachakraborti@gmail.com)

Md. Zonieed Hossain

e-mail: [zonieed.uap@gmail.com](mailto:zonieed.uap@gmail.com)

T. Ahamed

e-mail: [taukirahamed01@gmail.com](mailto:taukirahamed01@gmail.com)

Md. Abdul Hamid

e-mail: [ahamid@uap-bd.edu](mailto:ahamid@uap-bd.edu)

M. F. Mridha

e-mail: [firoz@uap-bd.edu](mailto:firoz@uap-bd.edu)

## 1 Introduction

Playing video games for entertainment to choosing profession as a gamer is definitely on the rise. But playing AAA blockbuster video games on a Computer is getting more expensive day by day. Game consoles (Play Station, Xbox, and Nintendo) are getting more optimization from developers while PC aren't. Hardware used in consoles are not that much powerful compared to PC, but playing games on it is way smoother than budget PCs. If optimization of games is possible for consoles, then it's also possible for PC. Game industries are not taking full care of optimization for their games. They are forcing people to buy expensive computer parts to play their games. They're making this industry for those people only who can get expensive computer parts. But it is not quite impossible to fully support their games in less expensive hardware. Applications for developing 3D video games already have their methods to optimize game assets. Developers just need to use it in the right way. Always there is a way to do more with optimization, it never ends. So here, our work shows how to optimize more 3D video game asset to run on less expensive hardware for indie developer to large industries. Our method for optimization is to use decimate modifier to less poly count and then we are removing the depending visualized players' unseen parts. Our method can be used for all game development software and computer games and mobile games.

The contribution of the paper is as follows:

- We propose a technique to optimize game assets with wiping out the concealed parts.
- We execute a method removing the invisible parts in high poly to low poly meshes.
- We have evaluated the performance of our proposed approach that optimizes average 50% of polygons.

The rest of the paper is organized as follows. In Sect. 2, we have discussed about the definition of optimization. In Sect. 3, we have discussed some works related to the proposed approach. We have discussed our proposed method with a complete description in Sect. 4. We have evaluated the performance of our proposed model in Sect. 5. Finally, in Sect. 6, we conclude our paper.

## 2 3D Mesh Optimization Definition

It generally means making things superior. **Improvement is possibly considered in the event that it creates precisely the same result.**

Calculating a very simple example:

$$x = (a + b) * (a + b) * (a + b)$$



Instead, calculating:

$$y = a + b, \text{ and setting } x = y * y * y$$

That would be an “optimization”, it only requires one addition rather than three for the exact same result.

### 3 Related Work

Just great quality amusements can hold their players, and this has turned into an imperative factor for any product diversion to succeed economically. At the end of the day, if amusement isn't of good quality, players can undoubtedly change to another diversion. Subsequently, it has turned out to be obligatory for the product diversion industry to endeavor to transform and adjust to the inclinations and requests of its players [1].

3D Modeling, Character Animation, Texture mapping, Collision location, Particle impacts, and Interaction code are engaged for streamlining of 3D versatile amusement advancement dependent on Unity. See Peng and Zhu [2] for an audit. They take a shot at Unity diversion advancement programming and for versatile recreations improvement.

In [3], to tackle the mesh optimization issue, they limit a vitality work that catches the contending wants of tight geometric fit and conservative portrayal. The trade-off between geometric fit and conservative portrayal is controlled by means of a client selectable parameter crep. An expansive estimation of crep demonstrates that an inadequate portrayal is to be emphatically favored over a thick one, generally to the detriment of corrupting the fit.

Lindstrom and Turk's objective is to exhibit a technique for enhancing the presence of an effectively rearranged model utilizing improvement that is guided by pictures [4]. This is a work enhancement technique that considers the geometry of a model as well as properties, for example, textures and textures normals [4]. This methodology fixes issues in an improved work that rearrangements strategies are unfeeling to, for example, splits between surface parts and item interpenetration [4].

In [5], they have led a progression of numerical analyses to figure out which of a few chose enhancement strategies are most reasonable for tackling the work shape quality streamlining issue where the majority of the vertices are at the same time repositioned to enhance normal quality. They thought about eight distinct solvers: six best in class solvers and two custom solvers they created [5]. They outline their discoveries as far as the techniques' strength, time to arrangement, adaptability to be utilized with an early halting standard, and versatility [5].

In [6], they made the mesh smoother. To make the improved mesh smoother, they pursued different strategies. They introduce a framework for triangle shape optimization and feature preserving smoothing of triangular meshes that is guided by the vertex Laplacian's, specifically, the uniformly weighted Laplacian and the discrete mean curvature normal. They provide different weighting schemes and demonstrate

the effectiveness of the framework on a number of detailed and highly irregular meshes [6].

According to [7], this paper shows a surface mesh enhancement technique, reasonable to acquire a geometric limited component mesh, given an underlying discretionary surface triangulation. A particular use of this system to the geometric mesh rearrangements is then laid out, which goes for diminishing the quantity of mesh entities while safeguarding the geometric estimate of the surface [7]. A few instances of surface cross sections planned for various application regions underline the productivity of the proposed approach.

Our proposal is to reduce the poly count without losing visual quality. So, gamers can enjoy the quality without compromising performance and also the game will be more playable in lower end hardware.

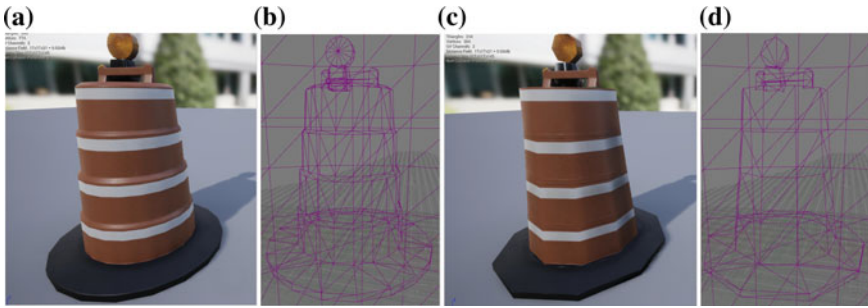
## 4 Proposed Method

Our proposal is to reduce poly count from 3D model and remove the parts of the scene. Removing parts are dependent on visualization. In a 3D video game, the player cannot see some parts that can be eliminated to save performance cost. A video game contained millions of assets. So, if we can eliminate unseen part of models, this can give huge amount of performance boost. To reduce poly count from 3D model, we generate an optimized method.

### 4.1 3D Asset Poly Complexity Reduction

Here, optimizing the raw mesh is done with decimate modifier. This decimate modifier grants you to reduce the vertices/triangles count of a mesh. It helps to convert a complex model with huge number of vertices and edges to a simple model with a smaller number of vertices and edges. It also helps to remove unwanted vertices and edges to increase the performance. The decimate modifier doesn't destroy the original mesh structure and it's an uncomplicated way to reduce the vertices and triangles of a mesh without any destruction of mesh.

In Fig. 1a, we can see a raw 3D model which is not optimized. Figure 1b is representing its wireframe view. Figure 1c is representing an optimized mesh of Fig. 1a. Figure 1d is the representation of wireframe of Fig. 1c. This method is discussed in [2] and [3]. Now it's clearly visualized that using decimate modifier decreases complexity but also losing quality.



**Fig. 1** Comparing raw mesh with optimized mesh using decimate modifier. **a** Raw mesh, **b** Wireframe view of raw mesh, **c** Optimization using decimate modifier, **d** Wireframe view of optimized mesh

### 4.2 3D Character Optimization

We are removing the invisible part in a game. We are considering a 3D model that has many 2D parts and every 2D part represents a two-dimensional array. Epitomizing our concept in an algorithm removes the unnecessary 2D part of any 3D model.

---

Algorithm 1. The procedure of initialize zero in invisible point.

---

Inputs:

- i: First cut point position number
- mesh[4][4]: An two dimensional array
  - {1, 2, 3, 4}
  - {5, 6, 7, 8}
  - {9, 10, 11, 12}
  - {13, 14, 15, 16}

- j: 0
- k: The value is less 1 than array size

Output:

- mesh[4][4]: Same structured as input but less complexity.
  - {1, 2, 3, 4}
  - {0, 0, 0, 0}
  - {0, 0, 0, 0}
  - {13, 14, 15, 16}

Procedure:

1. Start
2. Declare variable i and mesh[4][4]
3. Initialize the array values and the starting cut point position number in i
4. For i = first cut point position number to last cut point position number
5. Call function search\_cut(array[], array\_size, cut\_point\_position\_number)
6. Function search\_cut(array[], array\_size, cut\_point\_position\_number)

7. Declares variables j and k
8. Initialize variables  $j \leftarrow 0$  and  $k \leftarrow \text{array\_size}-1$
9. While j is less than array\_size AND k is greater than or equal 0
10. If array\_position\_number = cut\_point\_position\_number
  11. array\_position\_value  $\leftarrow 0$
  12. Display array\_position\_value
  13. Return 1
14. Else If array\_position\_number is greater than cut\_point\_position\_number
  15. Decrement of k
16. Else
  17. Increment of j
18. End While
19. Declare cut\_point\_position\_number is not found in array
20. Return 0
21. End Function
22. End For
23. Stop

---

The algorithm-1 illustrates a distinct concept of our work. Every step of algorithm-1 parades how our method works sequentially. First, we declare a two-dimensional array which is a two-dimensional part of a 3D character or a mesh and assign the starting cut point position number in variable i. Then, we have conducted a for loop from the starting cut point position number to last cut point position number. For every cut point position number, we put the array mesh[4][4], array size, and cut point position number in a function, called search\_out(array[], array\_size, cut\_point\_position\_number). In this function, declaring variables j equals 0 and k equals (array\_size – 1), then conducting a while loop j is less than array\_size and k is greater than or equal 0. In this loop, it will check that array position number is equal to cut\_point\_position\_number or not. If they both are equal, then assign the cut\_point\_positon\_value to 0.

Our primary target was to cut unnecessary parts from any kind of characters and meshes. In Fig. 2, we have presented an efficient flow to compare the unnecessary point with the structure of mesh. If part of a mesh is considered as an unnecessary point, we are removing these points with remaining the same structure of the mesh.

### 4.3 *Mathematical Formulation and Description*

Assuming the triangles set is  $T$ , vertices set is  $V$ ,

$$\text{Optimized\_mesh}(T, V) = \text{Raw\_mesh}(T, V) - \text{Cut\_point}(T, V).$$

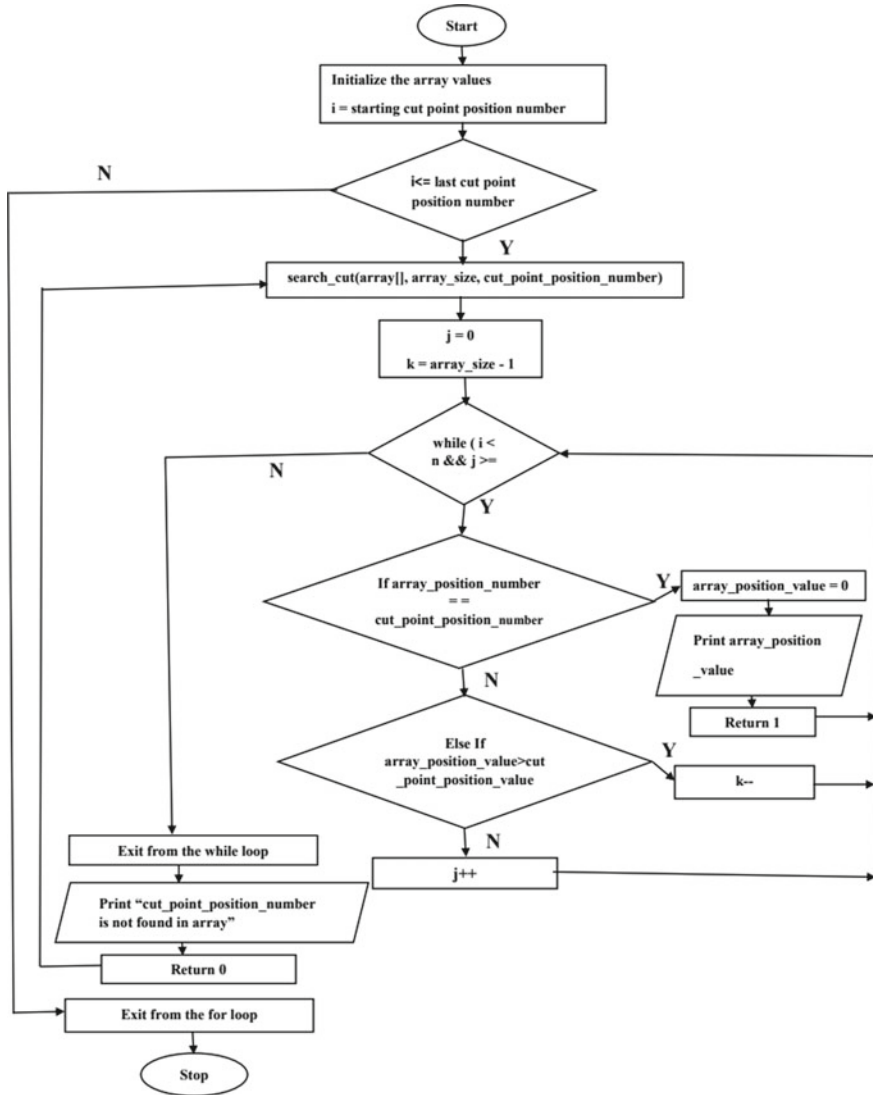


Fig. 2 Flowchart of proposed method

The simplified model as  $Optimized\_mesh(T, V)$  represents that the optimized 3D model contains  $T$  number of triangles and  $V$  number of vertices. The  $Raw\_mesh(T, V)$  represents that the raw 3D model contains  $T$  number of triangles and  $V$  number of vertices. The  $Cut\_point(T, V)$  contains  $T$  number of triangles and  $V$  number of vertices are going to be eliminated.

#### 4.4 Implementation of Our Proposed Method

Representing 2D part of rectangle 3D model called mesh in matrix

$$\begin{aligned} \text{mesh}[4][4] &= \{1, 2, 3, 4\} \\ &\quad \{5, 6, 7, 8\} \\ &\quad \{9, 10, 11, 12\} \\ &\quad \{13, 14, 15, 16\} \end{aligned}$$

These values can be varied mesh to mesh. Then, select the cutting point which is not visible to player in game but containing in mesh. Let's assume that in matrix mesh, the invisible points are

$$\begin{aligned} &\{5, 6, 7, 8\} \\ &\{9, 10, 11, 12\} \end{aligned}$$

Consider that  $i$  = first cut point position number

To eliminate these cutting points, repeat steps until

$i$  = last cut point position number

And in every step, we compare the mesh matrix value with cut point. If both are equal, then assign the array value zero.

The output result is as expected

$$\begin{aligned} \text{mesh}[4][4] &= \{1, 2, 3, 4\} \\ &\quad \{0, 0, 0, 0\} \\ &\quad \{0, 0, 0, 0\} \\ &\quad \{13, 14, 15, 16\} \end{aligned}$$

According to our proposed method, the result shows that the matrix mesh represents the structure as before with less complexity and less vertices and triangles count.

Now, if we represent a 2D part of a polygon and it contains these values

$$\begin{aligned} &\{1\} \\ &\quad \{2, 3, 4\} \\ &\quad \{5, 6, 7, 8\} \\ &\quad \{9, 10, 11, 12, 13\} \\ &\quad \{14, 15, 16, 17\} \\ &\quad \{18, 19, 20\} \\ &\quad \{21, 22\} \end{aligned}$$

Suppose that the imperceptible points which the player can't see in the game are

$$\begin{aligned} &\{9, 10, 11, 12, 13\} \\ &\{14, 15, \dots\} \\ &\{18, 19, \dots\} \\ &\{21, 22\} \end{aligned}$$

We will apply the same mechanism which is already described in the two examples before. The resulted output will be like as before with the same structure and less complexity and less vertices and triangles count.

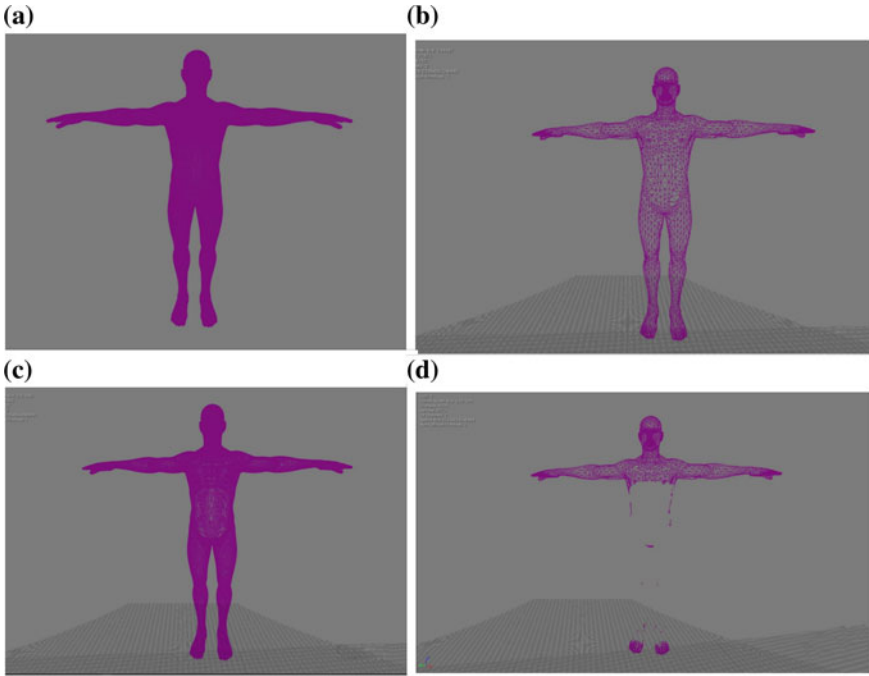
$$\begin{aligned} &\{1\} \\ &\{2, 3, 4\} \\ &\{5, 6, 7, 8\} \\ &\{0, 0, 0, 0, 0\} \\ &\{0, 0, 16, 17\} \\ &\{0, 0, 20\} \\ &\{0, 0\} \end{aligned}$$

According to these examples of different shapes, our method can work on all types of shapes in 3D model and it reduces the complexity of a game and reduces the number of vertices and triangles.

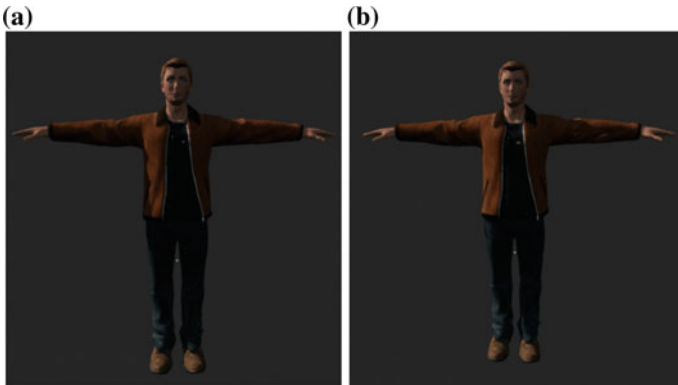
Figure 3a shows a raw model of a 3D Character which contains number of Triangles 468,542 and number of Vertices 240,248. Figure 3b shows an optimized model of Fig. 3a, which contains number of Triangles 11,571 and number of Vertices 6,752. We have used decimate modifier to reduce mesh complexity. There is no noticeable quality difference from standard camera distance as can be seen from the figure. But quality difference is noticeable from close range. Also, this model is not suitable for high-density display. Figure 3c shows 59.83% less complexity from Fig. 3a model, which contains number of Triangles 187,416 and number of Vertices 97,263. This model is good for high-density output but it requires more hardware resources.

Figure 3d shows a model with Triangles 9,233 and vertices 5,773 which doesn't look like a complete model but after applying cloths, there's no difference. Figure 4a represents Fig. 3a with added cloths and Fig. 4b represents Fig. 3d with added cloths. This is our proposed method to optimize 3D character in video games. Our method to optimize 3D character is depended on how developer wants to show their character in video games. As it depends on visualization, cut point of models should be handled with care.

This figure shows there is no visual difference between both models. However, cutting parts of characters depends on the clothing style. We have designed all these figures.



**Fig. 3** 3D character optimized with our proposed method. **a** shows a raw mesh, **b** shows same mesh with much less complexity, **c** same mesh less optimized but good for higher density display, **d** after applying our proposed method



**Fig. 4** Visual output of Fig. 3. **a** after applied cloths on raw mesh and **b** our proposed method applied mesh with cloths on



## 5 Result Analysis

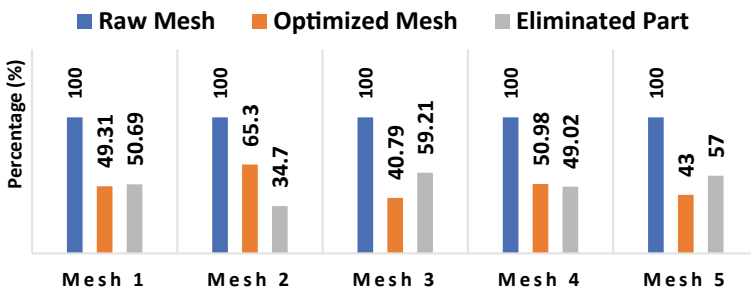
We have designed five raw meshes using Adobe Fuse CC and Blender 2.79b in Windows 10, then applied our optimization method.

In Table 1, first column shows mesh numbers, second column and third column represent number of triangles and vertices of raw meshes. Then, fourth and fifth column represent optimized version of those raw meshes. Finally, the last column shows the percentages of reducing the number of triangles and vertices from raw meshes. Our method depends on visual representation of meshes so output result may vary mesh to mesh. Our results show that we can reduce on an average of 50% from raw models without losing any quality. It is also possible that our method will work on very low poly to very high poly mesh.

Figure 5 represents raw mesh, optimized mesh, and eliminated part of each mesh in percentage. Here, we are comparing raw mesh with optimized mesh using our method and the part that we removed. All the data of this figure is taken from Table 1.

**Table 1** Performance of our optimization method

Mesh number	Number of triangles of raw mesh	Number of vertices of raw mesh	Number of triangles after applied our optimization method	Number of vertices after applied our optimization method	Percentage of optimization without losing quality (%)
1	281126	140994	136880	71249	50.69
2	1874176	949028	1218462	624951	34.7
3	181225	92254	74004	37553	59.21
4	562252	287270	286678	146411	49.02
5	51839	28178	22072	12333	57



**Fig. 5** Comparison between raw mesh and optimized mesh from Table 1

## 6 Conclusion

In this paper, we have discussed some challenges about character and mesh optimization of 3D video game. We have shown an effective way to optimize mesh without losing any quality and support in wide range of hardware. We have derived a technique that allows quality of mesh to be higher and complexity of mesh to lower. This method is also applicable for 3D animation, mobile games, or any other platforms where 3D contents are applied.

## References

1. Aleem, S., Capretz, L.F. & Ahmed, F. (2016, September). Critical success factors to improve the game development process from a developer's perspective. *Journal of Computer Science and Technology*, 31(5), 925–948.
2. Peng, H., & Zhu, K. (2014). Strategy research on the performance optimization of 3D mobile game development based on unity. *Journal of Chemical and Pharmaceutical Research*, 6(3), 785–791.
3. Hoppe, H., DeRose, T., Duchampy, T., McDonaldz, J., & Stuetzlez, W. (1994). *Mesh optimization* (pp. 19–25). Seattle, WA: University of Washington.
4. Lindstrom, P. & Turk, G. (2001). Image-driven mesh optimization. In *International Conference on Computer Graphics and Interactive Techniques*. Los Angeles, CAI, August 12–17, 2001.
5. Freitag, L., Knupp, P., Munson, T., & Shontz, S. (2002). *A comparison of optimization software for mesh shape quality improvement problems*. IMR.
6. Nealen, A., Igarashi, T., Sorkine, O., & Alexa, M. (2006). Laplacian mesh optimization. In *Proceedings of the 4th International Conference on Computer Graphics and Interactive Techniques in Australasia and Southeast Asia—GRAPHITE*. <https://doi.org/10.1145/1174429.1174494>.
7. Frey, P. J., & Borouchaki, H. (1998). Geometric surface mesh optimization. *Computing and Visualization in Science*, 1(3), 113–121. <https://doi.org/10.1007/s007910050011>.

# Image Watermarking Scheme Using Cuckoo Search Algorithm



Gaurav Dubey, Charu Agarwal, Santosh Kumar  
and Harivansh Pratap Singh

**Abstract** The work presented in this paper proposes a robust and optimized algorithm to be used in image watermarking. It makes use of cuckoo search algorithm (CSA) for optimization. A binary watermark is embedded within the host images in transform domain using a fitness function. The locations used for inserting the watermark bits are selected using CSA. The fitness function is a linear summation of similarity correlation coefficients  $SIM(W, W')$  obtained for four different operations performed on the watermarked image. The simulation reports indicate that the PSNR values are high enough. As a result, the watermarked images have high visual quality. The values of similarity correlation coefficient show that the scheme presented in this paper is also robust against the said operations. It can be concluded that this CSA-based scheme shows good result and also reports better results than other similar works.

**Keywords** Grayscale watermarking · Cuckoo optimization algorithm · Optimization · PSNR · Similarity correlation · Robustness

---

G. Dubey (✉) · S. Kumar · H. P. Singh

Department of Computer Science and Engineering, ABES Engineering College, Ghaziabad, India  
e-mail: [gdubey1977@gmail.com](mailto:gdubey1977@gmail.com)

S. Kumar

e-mail: [santoshg25@gmail.com](mailto:santoshg25@gmail.com)

H. P. Singh

e-mail: [harivansh.singh@abes.ac.in](mailto:harivansh.singh@abes.ac.in)

C. Agarwal

Department of Computer Science and Engineering, Ajay Kumar Garg Engineering College,  
Ghaziabad, India  
e-mail: [agarwalcharu2@rediffmail.com](mailto:agarwalcharu2@rediffmail.com)

© Springer Nature Singapore Pte Ltd. 2020

M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,

Lecture Notes in Networks and Systems 94,

[https://doi.org/10.1007/978-981-15-0694-9\\_61](https://doi.org/10.1007/978-981-15-0694-9_61)

## 1 Introduction

Image watermarking has now been established as a potent tool to check copyright violation of various digital media [1–3]. To this end, several algorithms have been proposed in the literature during the last 15 years or so. Two major criteria are to be mandatorily fulfilled by any image watermarking algorithms are high perceptual quality and high robustness values. As these two criteria are clashing with each other, so to optimize them, image watermarking methods are used. At present, the focus of developing an effective watermark scheme is converged to the optimized usage of soft computing tools to produce best results. Besides the aforesaid criteria, number of bits that can be embedded in an image is also crucial to this scheme. However, this issue is not given much importance and is taken as a constant, as the size of host signal is very large as compared to the watermark size. Several researchers worldwide have successfully used related soft computing tools to achieve these goals. A brief survey of these methods is given as under.

Shieh et al. [4] have developed a genetic algorithm (GA) based optimized image watermarking algorithm in the transform domain. They have proposed GA-based optimization of this process by considering the above discussed conflicting requirements. In this paper, they select a fitness function as given by Eq. 1.

$$f_c = \text{PSNR}_c + \sum_{h=1}^p (NC_{c,h} \cdot \lambda_{c,h}) \quad (1)$$

According to the author, their proposed scheme survives against various different image processing attacks discussed in the literature and also an improvement PSNR values by the application of their genetic algorithm (GA).

Wei et al. [5] also employed a GA to identify the best suitable positions (locations) to embed a binary watermark. They use a fitness function given by Eq. 2 which is a linear summation of similarity correlation coefficients computed during various iterations.

$$\text{Fitness} = \sum_{i=1}^4 \text{SIM}_i \quad (2)$$

They claim to have obtained signed images having a watermark invisible to human eyes. According to them, their proposed method can survive against different image processing operations.

Huang et al. [6] have developed a scheme using fuzzy theory to design the fitness function given by Eq. 3.

$$f_i = \text{PSNR}_i + \lambda \cdot \text{BCR}_i \quad (3)$$

In their work, they claim that by using fuzzy concept in conjunction with bacterial foraging, they obtained better results over existing implementations with GA.

Recently, Yang [7] has developed an optimization algorithm known as cuckoo search algorithm. In the present paper, we make use of cuckoo search algorithm to stabilize the trade-off between the conflicting requirements of watermarking. The fitness function used in our experiment is the same as given by Eq. (2). The proposed scheme successfully performs embedding, extraction of a binary watermark from grayscale images. The issue of robustness is also taken up in the experimental simulation by carrying out four image processing operations—Low-pass filter (radius = 0.1), Scaling (256-512-256), Gaussian noise addition  $N(0, 0.0)$ , and JPEG compression ( $Q = 5$ ). However, after successive each iteration of attacks, the watermark  $W'$  is extracted from the image and its  $SIM(W, W')$  is computed which is further used to reoptimize the fitness function given by Eq. 2.

## 2 Cuckoo Search Algorithm

Swarm-based cuckoo search (CS) algorithm was given by Xin-She Yang and Suash Deb in 2009 [7]. It is an optimization technique which is used already to optimize many applications [8]. It basically works on the concept of cuckoo bird that generally lay their eggs on other birds' nests and also they have high reproduction rate. Cuckoos have parasitic relationship, in which one species is benefited while another is harmed. Cuckoos eggs can be found in communal nests. To feed their own eggs and to protect them, they may remove host's nest eggs or babies.

Rules which describe the cuckoo search algorithm are as mentioned in [7]:

- (1) Each cuckoo randomly selects one nest to lay its one egg.
- (2) Select the nests to be considered in the next generation which consist of maximum good quality eggs

According to Yang [7], a solution is represented by an egg inside a nest and a cuckoo egg represents a new solution. Cuckoo search algorithm is developed to make use of new and higher objective value solutions with lower objective solutions.

A cuckoo  $i$  makes use of Lévy flight to generate a new solution  $x^{(t+1)}$ , which is given by Eq. 5.

$$x_i^{(t+1)} = x_i^t + \alpha \oplus Lévy(\lambda) \tag{5}$$

where  $\alpha$  should be greater than 0, denotes step size. Generally,  $\alpha$  is taken as one [7, 8]. Equation 5 determines next location of the cuckoo using two values: the present location of the cuckoo and probability of movement based on Lévy flight. Hadamard product ( $\oplus$ ) is performed in second term. Listing 1 gives the pseudo code of the CS.

**Listing 1:**

**Input:** Define fitness function  $f(y)$ ,  $y = [y_1, y_2, \dots, \dots, y_d]^T$

Initialize m host nest randomly, where each nest if given by  $y_p (p = 1, 2, \dots, m)$

**Begin**

**while** ( $n < \text{Maximum Generation}$ )

    Obtain randomly, a cuckoo by Lévy flight

    Calculate its fitness  $F_l$

    From available m host nests randomly (say, k) , select a nest

**If** ( $F_l > F_k$ ),

    New solution takes the place of nest k;

**End**

    From the available m host nest, poor quality nests ( $p_a$ ) are left;

    Retain good value solutions;

    Arrange solutions according to their fitness value and select the solution with highest fitness value

**End while**

    Post processing of the generated results

**End**

### 3 Proposed Scheme

In the proposed algorithm, to embed into host image, a one-bit image of size  $p \times q$  is used. Generally, the optimization tools which are used for this purpose operate on either of the two strategies—optimization of the embedding method [1, 8, 9] or to identify coefficients to be selected for embedding followed by actual embedding by using a generic formulation [4–6]. As mentioned earlier, we propose a watermark embedding and extraction scheme involving the cuckoo search algorithm. In this case, we use it to identify optimized locations for embedding the coefficients of the binary image used as original watermark ( $W$ ). The fitness function used in this algorithm is same as one given by Eq. 2. This is a linear summation of four similarity correlation values. Correlation values are computed between the embedded one-bit image ( $W$ ) and extracted one-bit image ( $W'$ ). Here,  $W'$  is the extracted watermark which is recovered from four different images obtained by executing four different attacks over signed images. The similarity correlation is computed by using the formulation given in Eq. 8.

$$SIM(W, W') = \frac{\sum_{i=1}^p \sum_{j=1}^q [W(i, j) \cdot W'(i, j)]}{\sum_{i=1}^p \sum_{j=1}^q \sqrt{W' \cdot W'}} \tag{8}$$

#### A. Watermark Embedding

For embedding, we consider an image  $X$  which has  $M$  rows and  $N$  columns and a binary watermark  $W$  of size  $(p \times q)$ . The algorithm used to embed the one-bit watermark image is listed in Listing 2.

Listing 2:

Step 1: Decompose the given image  $X$  into  $8 \times 8$  sized blocks. Convert each block into a frequency domain using discrete cosine transform. To each block, apply cuckoo search algorithm to coefficients 1–63 only. The 0th coefficient in the sequence being DC coefficient is excluded.

Step 2: To identify locations using cuckoo search algorithm do as follows:

- (a) Generate  $m$  cuckoos' randomly as the initial population, where each cuckoo is a vector of size  $c$ . Here,  $c$  is a vector of size  $(p \times q)$  which represents the watermark embedding locations in the host image. Then, from each  $8 \times 8$  block, only one location is selected for embedding
- (b) Embed the watermark bits (1/0) using Eq. 9 at locations selected by cuckoo search algorithm
- (c) Compute IDCT of the signed image to obtain signed image in the spatial domain. Obtain  $m$  such signed images separately
- (d) Apply four image processing attacks on each signed image one at a time. These are Low-pass filter (radius = 0.1), Scaling (256-512-256), Gaussian noise addition  $N(0, 0.0)$  and JPEG compression ( $Q = 5$ )
- (e) Use these four attacked images one at a time and thus extract the watermarks ( $W'$ s). Compare  $W$  and  $W'$  using  $SIM(W, W')$  for each of the four attacked signed images for  $m$  cuckoos
- (f) Find fitness of  $m$  cuckoos using Eq. 2
- (g) Move these  $m$  cuckoos as per the scheme listed in Listing 1.

Step 3: Select signed image corresponding to the cuckoo with the highest fitness value.

Equation 9 gives the formulation which is used in this scheme for modifying the host image in transform domain with one-bit watermark

$$V' = V * (1 + k * W) \quad (9)$$

where  $V$  denotes the coefficient of discrete cosine transform of the image given by cuckoo search algorithm,  $W$  is the one-bit image,  $k$  is the watermark strength, and  $V'$  is the discrete cosine transform coefficient of the watermarked image. In the present work, the value of  $k$  is taken as 0.4 after performing many simulations. The perceptual quality of watermarked images can be measured by PSNR.

## B. Watermark Extraction

The algorithm used to recover embedded watermark from the watermarked image is listed in Listing 3.

Listing 3:

Step 1: Compute block-wise DCT of original image and watermarked image

Step 2: Eq. 10 is used to recover the values of watermark image

$$W' = ((V'/V) - 1)/k \quad (10)$$

Step 3: Recover the watermark

Step 4: Compare  $W$  with  $W'$  using  $SIM(W, W')$

## 4 Experimental Report

Four 8-bit  $256 \times 256$  size images are used to evaluate the results of the proposed scheme. Namely, Boat, Baboon, Lena, and Pepper images are used as images to perform embedding operation. On the other hand, one-bit image of size  $32 \times 32$  is used as watermark. To conduct out simulation, we consider  $n = 20$  nests,  $\alpha = 1$ ,  $p_a = 0.5$ , ML (max iterations) = 10, and  $m$  (no. of cuckoos) = 10. The attacks used for computing fitness values are Low-pass filter (radius = 0.1), Scaling (256-512-256), Gaussian noise addition  $N(0, 0.0)$ , and JPEG compression ( $Q = 5$ ). Only two images are shown in this paper due to space constraints. The simulation results after carrying out attacks are given for all four images in Table 1. Figure 1a–d depicts the original and signed Boat and Baboon images. Figure 2a–c depicts the original and recovered watermarks, respectively. The  $SIM(W, W')$  value obtained out of recovered watermark is mentioned on top of Fig. 2b–c.

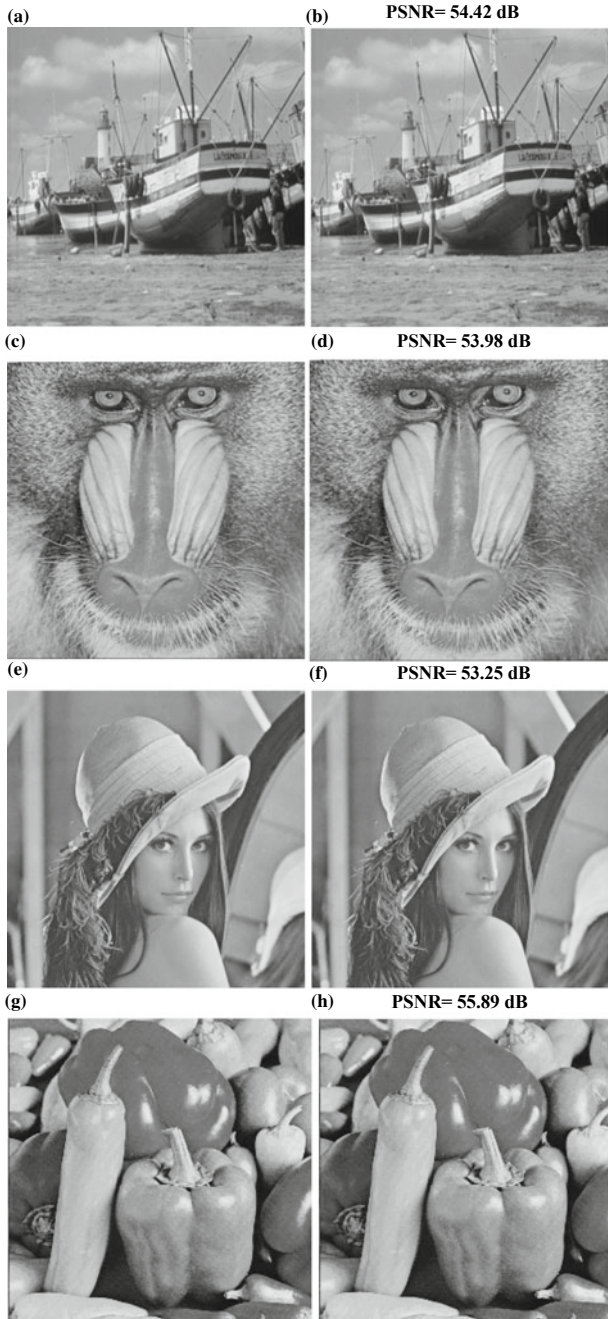
The PSNR and  $SIM(W, W')$  values obtained by the four host images are tabulated in Table 1. It also gives a comparison for these parameters as computed by Wei et al. [5].

In this table, SIM1, SIM2, SIM3, and SIM4, respectively, represent the similarity correlation parameter values for four image processing operations as described earlier. The values tabulated in above table shows that the reported values are higher than the other works done by [5]. The embedding and extraction are quite successful as indicated by high SIM values and perceptual parameter values. This is specifically true for all four images we use in our simulation and for all attacks barring the low-pass filter attack.

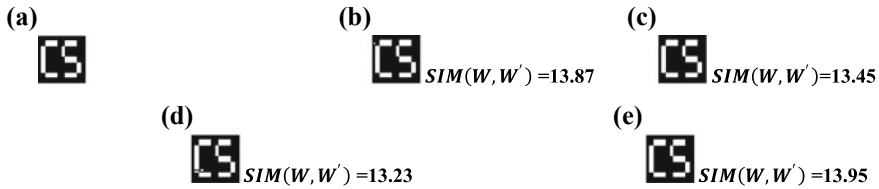
**Table 1** Comparison of PSNR and  $SIM(W, W')$  between our algorithm and one proposed by Wei et al. [5]

Image	Scheme	PSNR	SIM1	SIM2	SIM3	SIM4
Boat	Our work	54.42	6.782	8.872	11.342	10.342
	Ref. [5]	51.68	8.588	7.626	6.672	7.241
Baboon	Our work	53.98	6.385	8.451	11.981	11.874
	Ref. [5]	51.87	7.121	7.161	6.192	8.565
Lena	Our work	53.25	7.745	7.890	11.451	9.591
	Ref. [5]	51.55	7.579	7.289	7.715	8.247
Pepper	Our work	55.89	6.210	7.173	10.873	10.536
	Ref. [5]	51.86	7.936	7.158	7.726	7.403





**Fig. 1** Original image **a** Boat, **c** baboon, **e** Lena, **g** pepper, Signed image **b** Boat, **d** baboon, **e** Lena, and **h** pepper



**Fig. 2** **a** Original watermark, recovered watermark, **b** from Fig. 1b, **c** from Fig. 1d, **d** from Fig. 1f, and **e** from Fig. 1h

## 5 Conclusion

The proposed scheme gives the successful implementation of an optimized image embedding and extraction scheme performed using an optimization tool commonly known as cuckoo search algorithm. The proposed scheme identifies and selects the locations for embedding using cuckoo search algorithm. The fitness function is a linear summation of four similarity correlation values. Experimental values show that we have obtained high PSNR values. This indicates a high resemblance between the original image and signed image. During successive iterations of cuckoo search algorithm, four image processing operations are performed. PSNR and  $SIM(W, W')$  are computed after every successive iteration. It is found that the proposed scheme shows robustness for different image operations. We conclude that our scheme successfully optimizes the problem of image watermarking and performed better than other schemes based on evolutionary models of computation such as the one based on genetic algorithm (GA).

## References

1. Mishra, A., Agarwal, C., Sharma, A., & Bedi, P. (2014). Optimized gray-scale image watermarking using DWT–SVD and firefly algorithm. *Expert Systems with Applications*.
2. Agarwal, C., Mishra, A., & Sharma, A. (2015). A novel gray-scale image watermarking using hybrid fuzzy-BPN architecture. *Egyptian Informatics Journal*.
3. Agarwal, C., Mishra, A., & Sharma, A. (2011). Genetic algorithm-backpropagation network hybrid architecture for grayscale image watermarking in DCT domain. In *2011 Seventh International Conference on Intelligent Information Hiding and Multimedia Signal Processing*.
4. Shieh, C. S., Huang, H. C., Wang, F. H., & Pan, J. S. (2004). Genetic watermarking based on transform domain techniques. *Pattern Recognition*, 37(3), 556–565.
5. Wei, Z., Li, H., Dai, J., & Wang, S. (2006). Image watermarking based on genetic algorithm. In *IEEE International Conference on Multimedia and Expo, ICME* (pp. 1117–1120). Toronto, Ontario, Canada, July 9–12, 2006.
6. Huang, H. C., Chen, Y. H., & Abraham, A. (2010). Optimized watermarking using swarm-based bacterial foraging. *Journal of Information Hiding and Multimedia Signal Processing*, 1(1), 51–58.
7. Yang, X. S., & Deb, S. (2009). Cuckoo search via Lévy flights. In *Proceedings of World Congress on Nature & Biologically Inspired Computing (NaBIC 2009, India)* (pp. 210–214).

8. Mishra, A., & Agarwal, C. (2016). Toward optimal watermarking of grayscale images using the multiple scaling factor-based cuckoo search technique. In *Bio-inspired computation and applications in image processing* (pp. 131–155). Elsevier. <http://dx.doi.org/10.1016/B978-0-12-804536-7.00007-7>.
9. Huang, S., Zhang, W., Feng, W., & Yang, H. (2008). Blind watermarking scheme based on neural network. In *Proceeding of 7th World Congress on Intelligent Control and Automation (WCICA 2008)* (pp. 5985–5989).
10. Lou, D.-C., Hu, M.-C., & Liu, J.-L. (2008). Healthcare image watermarking scheme based on human visual model and back-propagation network. *Journal of C.C.I.T.*, 37(1), 151–162.
11. Agarwal, C., Mishra, A. & Sharma, A. (2011). Digital image watermarking in DCT domain using fuzzy inference system. In *Proceedings of 24th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE 2011)* (pp. 822–825). Niagara Falls, Ontario, Canada, May 8–11.

# A Survey of Latent Fingerprint Indexing and Segmentation Based Matching



Harivans Pratap Singh and Priti Dimri

**Abstract** Over the past few years, fingerprints have been considered the most sensitive and crucial identification basis for law enforcement agencies. In crime scene and forensics, recording of latent fingerprints from uneven and noisy surface is a difficult task and conventional algorithm fails in most of the times. A robust orientation field estimation algorithm is the need of the time to recognize the poor quality latent. To overcome the limitations of conventional algorithm, various techniques have been proposed in the last decade. In this paper, a comparative study has been done of state-of-the-art techniques with their advancements and limitations. Our proposal aims at effectively minimizing the difficulties faced to separate ridges and segmentation of latent images reducing search time and computational complexity while optimizing the system retrieval performance.

**Keywords** Latent · Indexing · Minutiae · Ridges · Skeleton · Coherence · Variation

## 1 Introduction

Fingerprint extraction has always been considered elementary to identify the suspects as it serves as an identification and evidence in crime-related cases because of its uniqueness and invariability with respect to a person regardless of passage of time. With technological advancement, there has been wide use of fingerprints tracing and identification in order to achieve maximum personal security. Fingerprints are employed to achieve security in case of cell phones, biometric door locks, attendance

---

H. P. Singh (✉)

Department of Computer Science and Engineering, UTU Dehradun, Sudhowala, India  
e-mail: [harivans.singh@abes.ac.in](mailto:harivans.singh@abes.ac.in)

ABES Engineering College Ghaziabad, Ghaziabad, India

P. Dimri

Department of Computer Science and Applications, G.B. Pant Engineering College, Ghurdauri, Uttarakhand, India  
e-mail: [2pdimri1@gmail.com](mailto:2pdimri1@gmail.com)

© Springer Nature Singapore Pte Ltd. 2020

M. L. Kolhe et al. (eds.), *Advances in Data and Information Sciences*,  
Lecture Notes in Networks and Systems 94,  
[https://doi.org/10.1007/978-981-15-0694-9\\_62](https://doi.org/10.1007/978-981-15-0694-9_62)

machines, and Indian aadhaar card which is among the largest biometrics usage around the world and one of the physiological characteristics of fingerprints is that they leave traces, i.e., once you leave your fingerprint over any object, they can be recorded later which is the main reason of recording them on crime locations in order to gather information about the criminal but the fingerprint traces are often in the latent form and fingerprint extraction from rolled and rough surfaces is a challenging task and there is every possibility of human error; therefore, rigorous study has been made and ample of methods have been invented for fingerprint identification using automation which minimizes errors. One of the proposed methods so far is indexing and fingerprint matching through segmentation where steps are involved from storing information of minutiae and generating skeleton structures in order to obtain a matching score. In this proposal, using the support vector machine (SVM), the latent fingerprints will be segregated through linear regression minutiae points and ridges. Features of latent images are further extracted using components such as gradient, ridge and image intensity, where gradients are studied by image edge detection technique, detecting accurate edges from the outline of the latent object through basic properties associated with an image like area, perimeter, and shape [1]. Ridges are set of two variables and curves whose points are local maxima in  $N-1$  dimensions the union of ridge sets and valley sets together form the connecting set of curves that intersect the critical points of the two-dimensional image which goes through the image intensity processing, the process includes adjusting the brightness and contrast with the image resulting in much easier visualization of the image where the intensity values are extracted from the multiple ROI.

Since fingerprint indexing is one of the most essential topics in the field of latent fingerprint matching techniques, it will allow us in prompt matching of the query against vast set of enrolled fingerprints called the search space without missing on explicit details.

## 2 Background Details and Related Work

Biometrics plays an important part in fingerprint indexing, majorly the two different components are verification and identification; similarity is measured between the patterns of two fingerprints, it also involves measurement of similar properties with reference to a single person whereas identification refers to matching of the patterns in the input finger with that of the database, for example, if database  $x$  has  $n$  number of total templates  $y$ , then  $x = \{y_0, y_1, \dots, y_n\}$  which means comparison will be performed to get the exact match of the fingerprint [2]. In defining the fingerprint biometrics, there are three parameters and matching which are pattern, minutiae points, pores and ridge shape [3].

1. Ridge Bifurcation: It is present in prints where one ridge splits into two ridges; it is found on every print irrespective of any fingerprint.

2. Pore: It gives the look of impression of holes in a ridge which are cavities and relatively very small but identifiable when enlarged properly.
3. Core: As you can see in Fig. 1, it is found in loops or whorls and usually seen where a ridge turns and runs back along itself with no ridge inside it and again it is very unlikely that the other person has one of these in the exact same position as you [4].
4. Delta: Mostly it is located in a whorl or loop print; it is formed when a triangular shape is obtained from ridges coming close together in different directions meeting at a point.
5. Ridge Ending: It is present in any fingerprints forming unique starting, most likely, anywhere in the print which ends suddenly forming the ending it gives individuality to a fingerprint.

The whorl, loop, arch kind of pattern cannot be related to an individual search but can be used for classification and indexing of fingerprint (Fig. 2).

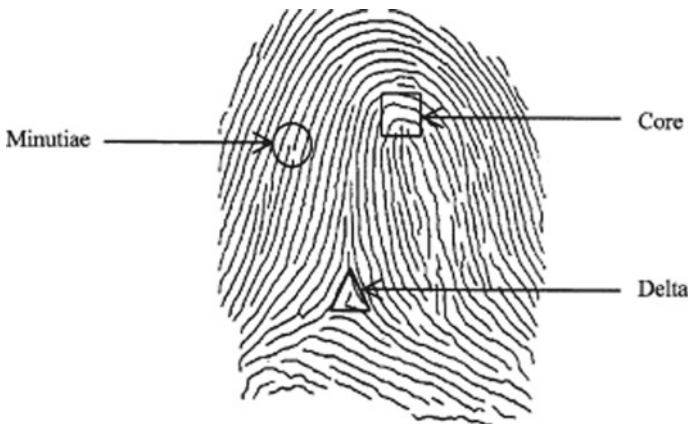
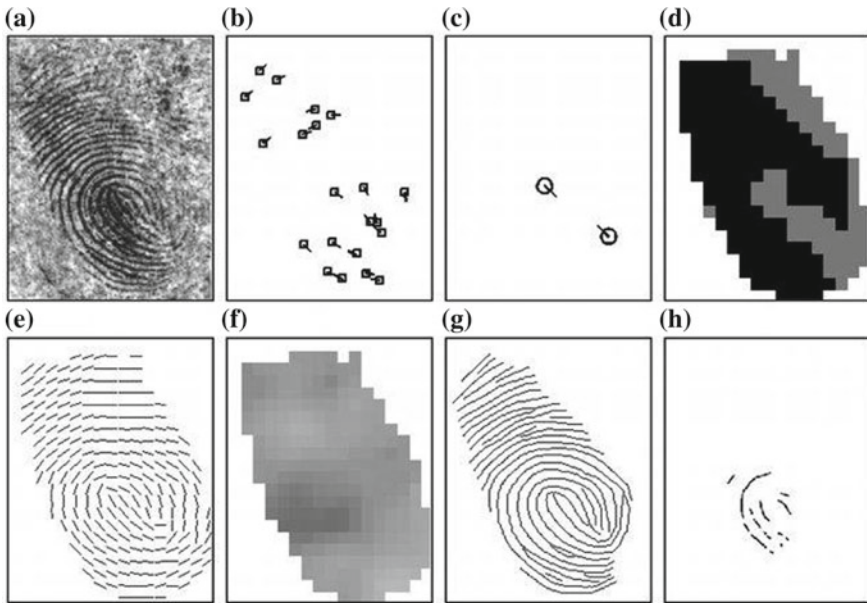


Fig. 1 Representation of minutia, core, and delta



Fig. 2 Arch, loop, and whorl



**Fig. 3** a Print gray-level image, b minutiae, c singularity points (cores), d ridge quality map, e ridge flow map, f ridge wavelength map, g skeletonised image, h incipient and dots ridges

Whereas features such as friction ridges, ridge deviations, ridge endings, lakes, islands, bifurcation, scars, incipient ridges. Various features in a fingerprint are as follows (Fig. 3):

Also, the features are further categorized into five groups based on their domain, named as follows:

- Features based on Saliency
- Features based on Intensity
- Features based on Gradient
- Features based on Ridge
- Features based on Quality.

These features are used to collect information through a fingerprint. To enable the lights out system [5], these features are to be extensively used in order to design such an algorithm that will give the desired output in place of a candidate list.

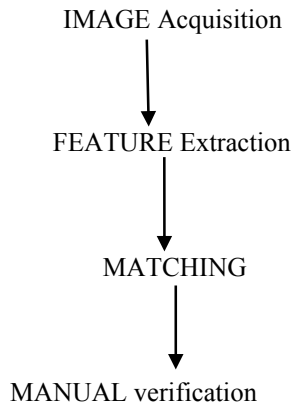
These features are used and studied to store information and perform matching. With the help of this matching score, the fingerprint is considered as a hit if the match is found, else a miss is considered [6]. The matching score is calculated with the running of algorithms and on the basis of matched minutiae and ridges. With the help of this matching score, a list is generated which is then used as a basis for latent examiner to match the fingerprints [7]. This is due to the current semi lights out system.

Hence, the existing fingerprint matching techniques focuses majorly on ridge patterns, minutiae-based approaches, and texture information.

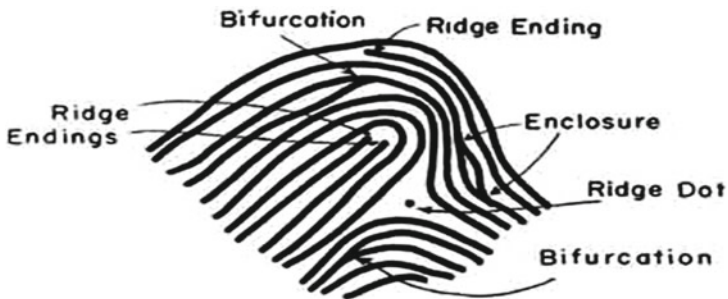
The local ridge orientation (LRO) is formed where the direction of the ridge flow structure is identified. Many studies have proven that ridge orientation is elementary in fingerprint image processing [8].

Ridge orientation is evaluated by LRO where every pixel located in the image and applications further includes filtering to improve or enhance images [9]. The ridge detection, when preprocessing and post -processing techniques, removes numbers of false minutiae and pattern analysis to extract classification types.

An automatic fingerprint identification (AFIS) based techniques consist of pre-processing stages as follows [10]:



While going through the preprocessing stages, various features of fingerprints are extracted such as an enclosure which is a collection of two bifurcations and a pair of ridge endings considered as a short ridge [11].



While going through these preprocessing stages, various features of fingerprints are extracted such as an enclosure which is a collection of two bifurcations and a pair of ridge endings considered as a short ridge [12].



As per the evaluation of latent fingerprint technologies (ELFT) conducted by NIST, the phase I result showed 80% accuracy among database of 10,000 rolled prints in identifying 100 latent images, whereas in the evaluation done by NIST, phase II identification accuracy was only 63.4% in the lights out identification after reviewing [13].

Through various studies, the common lights out system of fingerprint segmentation and matching involve the following process:

- Marking the segment of the fingerprint having best ridge information with the use of existing algorithms or through developing a new algorithm which can further segment the fingerprint more accurately and more quickly [14].
- After acquiring ridge skeleton map and ridge information, matching the segment against the fingerprints present in the database is done.
- The matching can be performed using any algorithm for that particular segment and not the whole latent fingerprint.
- After evaluating the matching score, in case of the presence of fingerprint in the database, further actions are performed.
- In case the fingerprints are not present in the database, different methods are then employed in order to reach to the suspect [15].

Cao and Jain [16] have proposed a Con Net-based fingerprint indexing approach where the performance of the stated indexing algorithm on two rolled fingerprint databases outperforms the old state of art indexing algorithms where the speed of fingerprint alignment is improved and thus improving the indexing accuracy.

As Sherlock and Monro [17] have presented simple model of fingerprint where the fingerprint topology and with regards to local ridge orientation, can be of practical use in the 2D interpolation with sampled LRO of real fingerprint.

Systematic methods with the computation of the singular points of fingerprints were proposed by Bazen [18] has developed a new PCA-based method for estimation which exemplifies consistency binary decisions and can be implemented very efficiently.

Enzhu Jinaping Yu [19] has proposed a scheme for systematic estimation of fingerprint ridge orientation and the correctness of orientation which is done by machine learning using neural network indicates the quality of the block and thus estimates the ridges which are not useful for segmentation and correctness of falsely estimated orientation.

Liu and Jiang [20] proposes a computational rational point where orientation for all types of fingerprints can be used for translation and rational alignment.

Zhou [21] member of IEEE, has calculated gradient-based algorithm and the filer bank based approach on a comparability stage shows that they are inaccurate and computationally expensive where the coarse filed computed using the gradient-based algorithm in which the error from the noise can be eliminated using weighted approximation.

Hong, Wan, and Jain [22] work proposed where image enhancement is done by improving both the goodness index and verification performance. A global model

of the ridges and furrows that can be made with the partial valid regions that can be utilized to correct the errors in the estimated images which will improve enhancement.

Jianjiang and Feng [23] have used approach in which they used a method FpVTE to match fingerprints. This evaluation showed that most commercial fingerprint matchers achieved an impressive rank 1 identification rate and proposed a skeleton algorithm which works on these pillars or basic keynotes:

1. The singularity of minutiae of two fingerprints is calculated
2. For every five most similar minutiae pairs, procedures are completed to establish the relation between skeletons of two fingerprints and a matching score is computed
3. The associated skeletons of the initial minutiae pair are assumed to be matched and used as a reference.
4. Skeletons adjacent to reference skeleton pairs are united according to reference skeleton pairs and then coordinated. This is iteratively performed until no more skeletons can be matched.
5. A skeleton matching score is computed.

The CDEFFS [24] was used during this research which documents most of the extended features such as virtual reference point, skeletonized image, ridge flow map, ridge quality map, crease, dot, incipient ridge, and pore.

The matching algorithm used in this approach was skeleton matching algorithm. This matching algorithm is stated as

- Select the base-paired minutiae (bpm) out of all the matched minutiae pairs. This is the most reliable minutiae pair.
- Screening of minutiae pairs then takes place and all those minutiae pairs are removed which are inconsistent with the bpm.
- The next step includes modifying the two skeleton images to make them more familiar for comparison.
- The final step includes matching of skeleton points supervised by matched minutiae and skeleton points incrementally.

Jain and Feng proposed the algorithm which works in this order:

1. Singularity between minutiae of two fingerprints is recorded
2. Five most similar minutiae points are recorded and for each pair, steps are performed in order to establish a relationship between skeletons of two fingerprints and on the basis of this relationship, matching score is computed.
3. The allied skeleton of the initial minutiae pairs are unspecified to be matched and are used as a reference for other minutiae pairs.
4. Skeletons adjacent to reference skeleton pairs are aligned according to reference skeleton pairs and then matched [25]. This is performed until there is no last skeleton and furthermore there are no more skeletons to be matched.
5. Finally, a skeleton matching score is thus calculated.

### 3 Proposed Approach

Our paper aims at comparing the state of art techniques and the indexing dynamics where the segmentation of fingerprint biometrics can be understood and thus the fingerprint topology can be improved. Most fingerprints when extracted are in rough and very noisy form so extraction of the same in the same form becomes very challenging and gathering maximum ridge information sometimes becomes difficult. Therefore, obtaining a rotational dimensional image through SVM model will be our approach. A thorough comparison and in detail methodology of the preprocessing and the post-processing stages will be studied.

### 4 Conclusions

Through various approaches, it has been clear that fingerprint matching system has drastically improved and has improved to being more accurate and efficient. The features and biometrics of fingerprint indexing have been thoroughly studied which has made the fingerprint segmentation more understanding to reach a light out system. Through various approaches, it is proven that minutiae and ridge orientation filed study has made fingerprint indexing more efficient.

### References

1. Arora, S., Cao, K., & Jain, A. K. (2014). Latent fingerprint matching: Performance gain via feedback from exemplar prints.
2. Cao, K., & Jain, A. K. *Fingerprint indexing and matching an integrated approach*. Michigan.
3. Sherlock, B. G., & Monro, D. M. (1992). *A model for interpreting fingerprint*. USA.
4. Brazen, A. M. & Gerez, S. H. (2002). *Systematic methods for the computation of the direction fields and singular points of fingerprints*. IEEE.
5. Zhu, E., & Yen, J. (2006). *Systematic method for fingerprint ridge orientation estimation and image processing*. China: Elsevier.
6. Liu, M., & Jiang, X. (2004). *Fingerprint reference point detection*. Singapore: EEE Nanyang Technological University.
7. Zhou, J. (2003). *A model-based method for the computation of fingerprints' orientation field*. China: IEEE.
8. Hong, L. (1998). *Fingerprint image enhancement: Algorithm and performance evaluation*. Michigan.
9. Jain, A. K., & Feng, J. (2011) *Latent fingerprint matching*. IEEE.
10. Karimi-Ashtiani, S., & Jay Kuo, C.-C. (2008). A robust technique for latent fingerprint image segmentation and enhancement. In *2008 15th IEEE International Conference on Image Processing*. IEEE.
11. Ruangsakul, P., et al. (2015). Latent fingerprints segmentation based on rearranged fourier subbands. In *2015 International Conference on Biometrics (ICB)*. IEEE.
12. Zhang, J., Lai, R., & Jay Kuo, C.-C. (2013). Adaptive directional total-variation model for latent fingerprint segmentation. *IEEE Transactions on Information Forensics and Security*, 8(8), 1261–1273.

13. Li, B., et al. (2011). Surface wrinkling patterns on a core-shell soft sphere. *Physical Review Letters*, 106(23), 234301.
14. Goswami, G., et al. (2013). On RGB-D face recognition using Kinect. In *2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*. IEEE.
15. Fu, F. (2005). Structural behavior and design methods of tensegrity domes. *Journal of Constructional Steel Research*, 61(1), 23–35.
16. Liang, X., Bishnu, A., & Asano, T. (2007). A robust fingerprint indexing scheme using minutia neighborhood structure and low-order Delaunay triangles. *IEEE Transactions on Information Forensics and Security*, 2(4), 721–733.
17. Moses, K. (2009). Automatic fingerprint identification systems (afis). In *Fingerprint source-book, international association for Identification*. Washington, DC: National Institute of Justice. <http://www.ncjrs.gov/pdffiles1/nij/225326.pdf>.
18. Paulino, A. A., Feng, J., & Jain, A. K. (2013). Latent fingerprint matching using descriptor-based hough transform. *IEEE Transactions on Information Forensics and Security*, 8(1), 31–45.
19. Jain, A. K., & Feng, J. (2011). Latent fingerprint matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(1), 88–100.
20. Yoon, S., Feng, J., & Jain, A. K. (2010). On latent fingerprint enhancement. In *Biometric technology for human identification VII* (Vol. 7667). International Society for Optics and Photonics.
21. Zhao, Q., Feng, J., & Jain, A. K. (2010). Latent fingerprint matching: Utility of level 3 features. *MSU Technical Report*, 8, 1–30.
22. Sankaran, A., et al. (2011). On matching latent to latent fingerprints. In *2011 International Joint Conference on Biometrics (IJCB)*. IEEE.
23. Sankaran, A., Jain, A., Vashisth, T., Vatsa, M., & Singh, R. (2017). A research paper on *Adaptive latent fingerprint segmentation using feature selection and random decision forest classification*. Elsevier.
24. Said, A., & Peralman, W. A. (1996). An image multiresolution representation for lossless and lossy image compression. *IEEE Transaction on Image Processing*, 5, 1303–1310.
25. Xiong, Z., Ramchandran, K., & Orchard, M. T. (1998). Wavelet packet image coding using space-frequency quantization. *IEEE Transaction on Image Processing*, 7, 892–898.

# Author Index

## A

Abhimanyu Kumar Patro, K., 81  
Acharya, Bibhudendra, 23, 67, 81, 179  
Adwani, Jyoti, 293  
Agarwal, Charu, 667  
Agarwal, Diwakar, 643  
Agarwal, Vanita, 237, 293  
Aggarwal, R. K., 425  
Aghav, Jagannath V., 283  
Agrawal, Diwakar, 533  
Ahamad, Mohd Vasim, 135  
Ahamed, Taukir, 655  
Ahmad, Ausaf, 119  
Ahmad, Riaz, 35, 119  
Aich, Payal, 497  
Alam, Bashir, 171  
Alam, Shadab, 35, 119  
Ally, Said, 381  
Ankith, K. U., 361

## B

Bagi, Randheer, 601, 613  
Bansal, Atul, 533, 643  
Bharadwaj, Skanda N., 361

## C

Chakraborti, Sumittra, 655  
Chaturvedi, Ankur, 433  
Chaturvedi, D. K., 317, 329  
Chaudhary, Akash Singh, 317, 329

## D

Deshpande, Amruta S., 217

Didwania, Bharat, 127  
Dimri, Priti, 677  
Doriya, Rajesh, 555  
Dr. Anil Sharma, 57  
Dubey, Gaurav, 351, 667  
Dutta, Tamina, 127  
Dutta, Tanima, 403, 591, 601, 613  
Dwivedi, Rajendra Kumar, 97, 441

## F

Farooq, Omar, 147, 545  
Fatima, Ghania, 545

## G

Goel, Pragati, 271  
Goel, Rati, 623  
Goel, Vikas, 271  
Gourisaria, Mahendra Kumar, 465  
Goyal, Vishal, 13  
Gupta, Amit Kumar, 271  
Gupta, Ashish, 127  
Gupta, Deepa, 497  
Gupta, Devbrat, 13  
Gupta, Hari Prabhat, 127, 403, 591, 601, 613  
Gupta, Sarthak, 47  
Gupta, Shelley, 477

## H

Hasan, Ragib, 655  
Hashmi, Anam, 147

## I

Imran, Mohd, 135

Isha, 317, 329  
 Ismail, Wan Khairuzzaman Wan, 521

**J**

Jain, Anmol, 623  
 Jamal, Tasleem, 135

**K**

Kamble, Bhawana, 555  
 Kaur, Gursimran, 511  
 Kaur, Jagreeti, 487  
 Khan, Bilal Alam, 147  
 Khan, Mohd Aamir, 577  
 Khatter, Ashish, 567  
 Kumar, Ajitesh, 307  
 Kumar, Ankit, 425  
 Kumar, Arvind, 395  
 Kumari, Mona, 307  
 Kumari, Nikita, 97  
 Kumar, Ishan, 343  
 Kumar, Jitendra, 13  
 Kumar, Rakesh, 97, 441  
 Kumar, Sanjay, 3  
 Kumar, Santosh, 351, 667  
 Kumar, Sumit, 487  
 Kumar, Sunil, 487  
 Kumar, Yogesh, 521

**L**

Luthra, Snigdha, 511

**M**

Malhotra, Virain, 47  
 Manuja, Prashant, 343  
 Md. Abdul Hamid, 655  
 Md. Zonieed Hossain, 655  
 Mishra, Rahul, 127  
 Mishra, Ruchi, 257  
 Mridha, M. F., 191, 655

**N**

Nagwani, Naresh Kumar, 3  
 Namdeo, Basant, 413  
 Nayak, Sinkon, 465  
 Negi, Sarita, 203  
 Nigam, Nitika, 403

**P**

Pal, Om, 171

Pandey, Manjusha, 455, 465  
 Pandey, Pranav, 283  
 Panwar, Neelam, 203  
 Patil, Rajendrakumar A., 237, 293  
 Patil, Sanjaykumar L., 217  
 Patro, K. Abhimanyu Kumar, 23, 67  
 Pavan Kumar, K., 67  
 Pradhan, Rahul, 433  
 Prasanth Jagapathi Babu, M., 67, 81  
 Prem Prakash, V., 247

**R**

Rafiq, Jahir Ibna, 191  
 Rajendra, A. B., 361  
 Ranjan, Jayanthi, 477  
 Rautaray, Siddharth Swarup, 455, 465  
 Rauthan, Man Mohan Singh, 203  
 Raza, Zahid, 109  
 Reddy, Nitya, 567  
 Rehman, Shabnum, 57

**S**

Sabharwal, Munish, 521  
 Sahoo, Bhaswati, 455  
 Sahu, Satya Prakash, 555  
 Saran, Munish, 441  
 Sharma, Diksha, 3  
 Sharma, Dilip Kumar, 433  
 Sharma, Harish, 257  
 Sharma, Nirmala, 257  
 Sharma, V. K., 23  
 Shreyas, V., 361  
 Shrivastava, Nivedita, 179  
 Shuaib, Mohammed, 35, 119  
 Siddiqui, Nazish, 135  
 Siddiqui, Shams Tabrez, 35, 119  
 Singh, Archana, 477, 487  
 Singh, Deepak Kumar, 227  
 Singh, Dilbag, 511  
 Singh, Gaurav, 127  
 Singh, Gurwinder, 57  
 Singh, Harivans Pratap, 677  
 Singh, Harivansh Pratap, 667  
 Singh, Juginder Pal, 577  
 Singh, Krishan Veer, 109  
 Singh, Shailendra Narayan, 47  
 Singh, Shailendra Pratap, 227  
 Singh, Shikha, 545  
 Sinha, Ankita, 455  
 Sinha, Nishant, 395  
 Soni, Aishwarya, 591

Soni, Shashwat, [23](#)  
Soni, Yashpal, [343](#)  
Sravanthi, Dasari, [81](#)  
Srinidhi, S., [361](#)  
Srivastava, Pallavi, [533](#)  
Subha, T., [157](#)  
Sudan, Amit, [521](#)  
Suman, Ugrasen, [413](#)

**T**

Thakur, Anita, [567](#)  
Tiwari, Shailesh, [351](#)  
Tripathi, Aprna, [433](#)

**V**

Vaisla, Kunwar Singh, [203](#)

Vashistha, Piyush, [577](#)  
Vats, Avaneesh Kumar, [373](#)  
Venugopalan, Manju, [497](#)

**W**

Wankhede, Nagsen, [373](#)

**Y**

Yadav, Mahima, [247](#)  
Yadav, Narendra Singh, [343](#)

**Z**

Zaman, Wahid Uz, [191](#)